

FIGURE 1

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGGAGGCACAGGTGGCCCCCACCACCCGGAGG
AGCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGA
AGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCT
TCTGGTGTTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTTAGGGTGTGTGCT
GTCCCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCC
TCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATTTATAGGACCGCCTAC
CGCCGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAG
GACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAG
GGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAG
TCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGC
CGGCAGTTACTGGTGCCAGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTG
TGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAG
GAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCT
GGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCC
TCCTGGTGCACCTCCTTCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTTC
CTGGAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGG
CTGGACTGAGCCCCCTCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTC
CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTTTTCTCCTC
CCCTTCCCTCGGGAGGGTCCCCAGACCCTGGCATGGGATGGGCTGGGATTTTTTTTGTGAAT
CCACCCCTGGCTACCCCCACCCTGGTTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCA
GCTGAGGGAAGGTACGAGTTCCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCC
CGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACATAAAAATGAAA
CGTGAAGGGCGGCCGCGACTCT
AGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGT
TACAAAT

FIGURE 2

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVASGVGSGRHRARLPARPLGCVLSRAHGDPV
SESFVQRVYQPFLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC
GAAICQPPCRNGGSCVQPGRCRCPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCW
EGHSLSADGTLCVPKGPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLAS
QALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 93-97, 270-274

N-myristoylation sites.

amino acids 19-25, 78-84, 97-103, 100-106, 103-109, 157-163,
191-197, 265-271

Amidation site.

amino acids 26-30

Aspartic acid and asparagine hydroxylation site.

amino acids 152-164

Cell attachment sequence.

amino acids 130-133

EGF-like domain cysteine pattern signature.

amino acids 123-135

FIGURE 3

CGCTCGCCCCGTCGCCCCTCGCCTCCCCGCAGAGTCCCCTCGCGGCAGCAGATGTGTGTGGG
GTCAGCCCACGGCGGGGACTATGGGTGAAATTCGCGGCGCTCACGCACTACTGGCCCCCTGATC
CGGTTCTTGGTGCCCCCTGGGCATCACCAACATAGCCATCGACTTCGGGGAGCAGGCCTTGAA
CCGGGGCATTGCTGCTGTCAAGGAGGATGCAGTCGAGATGCTGGCCAGCTACGGGGCTGGCGT
ACTCCCTCATGAAGTTCTTTCACGGGTCCCATGAGTGACTTCAAAAATGTGGGGCCTGGTGTTT
GTGAACAGCAAGAGAGACAGGACCAAAGCCGTCCTGTGTATGGTGGTGGCAGGGGCCATCGC
TGCCGTCTTTTCACACACTGATAGCTTATAGTGATTTAGGATACTACATTATCAATAAACTGC
ACCATGTGGACGAGTCGGTGGGGAGCAAGACGAGAAGGGCCTTCCTGTACCTCGCCGCCCTTT
CCTTTTCATGGACGCAATGGCATGGACCCATGCTGGCATTCTCTTAAAACACAAATACAGTTT
CCTGGTGGGATGTGCCTCAATCTCAGATGTCATAGCTCAGGTTGTTTTTGTAGCCATTTTGC
TTCACAGTCACCTGGAATGCCGGGAGCCCCCTGCTCATCCCGATCCTCTCCTTGTACATGGGC
GCACTTGTGCGCTGCACCACCCTGTGCCTGGGCTACTACAAGAACATTACAGCATCATCCC
TGACAGAAGTGGCCCCGAGCTGGGGGGAGATGCAACAATAAGAAAGATGCTGAGCTTCTGGT
GGCCTTTGGCTCTAATTCTGGCCACACAGAGAATCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTCACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTTCG
ACAAGAATAACCCAGCAACAAACTGGTGAGCACGAGCAACACAGTCACGGCAGCCACATC
AAGAAGTTTACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTTCTGTATGTTTTGGAC
ACCCAACGTGTCTGAGAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCTTTTGGGATCTTCTCCTTCTTCCCAGTTCCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTCTTGGCCCCAGCTCTGTGCTGCG
GATCATCGTCCTCATCGCCAGCCTCGTGGTCTACCTACCTGGGGGTGCACGGTGCAGCCC
TGGGCGTGGGCTCCCTCCTGGCGGGCTTTGTGGGAGAATCCACCATGGTTCGCCATCGCTGCG
TGCTATGTCTACCGGAAGCAGAAAAAGAAGATGGAGAATGAGTCGGCCACGGAGGGGGAAGA
CTCTGCCATGACAGACATGCCTCCGACAGAGGAGGTGACAGACATCGTGGAATGAGAGAGG
AGAATGAATAAGGCACGGGACGCCATGGGCACTGCAGGGACGGTCAGTCAGGATGACACTTC
GGCATCATCTCTTCCCTCTCCCATCGTATTTTGTTCCTTTTTTTTGTTTTGTTTTGGTAAT
GAAAGAGGCCTTGATTTAAAGGTTTTCTGTCAATTCTCTAGCATACTGGGTATGCTCACACT
GACGGGGGGACCTAGTGAATGGTCTTTACTGTTGCTATGTAAAAACAAACGAAACAACTGAC
TTCATACCCCTGCCTCACGAAAACCCAAAAGACACAGCTGCCTCACGGTTGACGTTGTGTCC
TCCTCCCCTGGACAATCTCCTCTTGGAACCAAAGGACTGCAGCTGTGCCATCGCGCCTCGGT
CACCTTGCACAGCAGGCCACAGACTCTCCTGTCCCCCTTCATCGCTCTTAAGAATCAACAGG
TAAAACTCGGCTTCCTTTGATTTGCTTCCAGTCACATGGCCGTACAAAGAGATGGAGCCC
CGGTGGCCTCTTAAATTTCCCTTCTGCCACGGAGTTCGAAACCATCTACTCCACACATGCAG
GAGGCGGGTGGCACGCTGCAGCCCGGAGTCCCGTTTCACTGAGGAACGGAGACCTGTGAC
CACAGCAGGCTGACAGATGGACAGAATCTCCCGTAGAAAGGTTTGGTTTGAAATGCCCCGGG
GGCAGCAAACTGACATGGTTGAATGATAGCATTTCACTCTGCGTTCTCCTAGATCTGAGCAA
GCTGTCAGTTCTCACCCCCACCGTGTATATACATGAGCTAACTTTTTTAAATTGTACAAAA
GCGCATCTCCAGATTCCAGACCCTGCCGCATGACTTTTCTGAAGGCTTGCTTTTCCCTCGC
CTTTCTGAAGGTCGCATTAGAGCGAGTCACATGGAGCATCCTAACTTTGCATTTTATGTTTT
TACAGTGAAGTGAAGCTTTAAGTCTCATCCAGCATTCTAATGCCAGGTTGCTGTAGGGTAAC
TTTTGAAGTAGATATATTACCTGGTTCTGCTATCCTTAGTCATAACTCTGCGGTACAGGTAA
TTGAGAATGTACTACGGTACTTCCCTCCACACCATAACGATAAAGCAAGACATTTTATAACG
ATACCAGAGTCACTATGTGGTCCCTCCCTGAAATAACGCATTTCGAAATCCATGCAGTGCAGTA
TATTTTTCTAAGTTTTTGAAAGCAGGTTTTTTCTTTTAAAAAATTATAGACACGGTTCCT
AAATTGATTTAGTCAGAATTCTAGACTGAAAGAACCTAAACAAAAAATATTTTAAAGATA
TAAATATATGCTGTATATGTTATGTAATTTATTTTAGGCTATAATACATTTCTATTTTCGC
ATTTTCAATAAAATGTCTCTAATAACAAAAA

FIGURE 4

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF
TGPMSDFKNVGLVFNNSKRDRTKAVLCMVVAGAIAAVFHTLIAYSDLGYIIINKLHHVDES
GSKTRRAFLYLAAPPFMDAMAWTHAGILLKHKYSFLVGCASISDVIAQVVFVAILLHSHLEC
REPLLIPILSLYMGALVRCTTLCGLGYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL
ATQORISRPIVNLFVSRDLGGSSAATEAVAILTATYPVGHMPYGWLTEIRAVYPAFDKNPSN
KLVSTSNVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFaelcVPLR
IFSFFPVPVTVRAHLTGWLMTLKKTFVLAPSSVLRIIVLIASLVVLPYLGVHGATLGVGSL
AGFVGESTMVAIAACYVYRKQKKMENESATEGEDSAMTDMPTTEEVTDIVEMREENE

Transmembrane domains:

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,
408-423, 431-445

FIGURE 5

CCTGACAGAAAGTGCCCCGGAGCTGGGGGAGATNCAACATTAAGAAGATGCTGAGCTTCTGGT
GCCNTTTGGCTCTAATTCTGGCCACACAGAGAANCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTCACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTTCG
ACAAGAATAACCCCAGCAACAACTGGTGAGCACGAGCAACACAGTCACGGCGGCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGNGAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCTTTGCGGATCTTCTCCTTCTTCCCAGTTCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTC

Parameter	1990-1991		1991-1992		1992-1993		1993-1994		1994-1995		1995-1996		1996-1997		1997-1998		1998-1999		1999-2000		2000-2001		2001-2002		2002-2003		2003-2004		2004-2005		2005-2006		2006-2007		2007-2008		2008-2009		2009-2010		2010-2011		2011-2012		2012-2013		2013-2014		2014-2015		2015-2016		2016-2017		2017-2018		2018-2019		2019-2020		2020-2021		2021-2022		2022-2023		2023-2024		2024-2025		2025-2026		2026-2027		2027-2028		2028-2029		2029-2030		2030-2031		2031-2032		2032-2033		2033-2034		2034-2035		2035-2036		2036-2037		2037-2038		2038-2039		2039-2040		2040-2041		2041-2042		2042-2043		2043-2044		2044-2045		2045-2046		2046-2047		2047-2048		2048-2049		2049-2050		2050-2051		2051-2052		2052-2053		2053-2054		2054-2055		2055-2056		2056-2057		2057-2058		2058-2059		2059-2060		2060-2061		2061-2062		2062-2063		2063-2064		2064-2065		2065-2066		2066-2067		2067-2068		2068-2069		2069-2070		2070-2071		2071-2072		2072-2073		2073-2074		2074-2075		2075-2076		2076-2077		2077-2078		2078-2079		2079-2080		2080-2081		2081-2082		2082-2083		2083-2084		2084-2085		2085-2086		2086-2087		2087-2088		2088-2089		2089-2090		2090-2091		2091-2092		2092-2093		2093-2094		2094-2095		2095-2096		2096-2097		2097-2098		2098-2099		2099-2100		2100-2101		2101-2102		2102-2103		2103-2104		2104-2105		2105-2106		2106-2107		2107-2108		2108-2109		2109-2110		2110-2111		2111-2112		2112-2113		2113-2114		2114-2115		2115-2116		2116-2117		2117-2118		2118-2119		2119-2120		2120-2121		2121-2122		2122-2123		2123-2124		2124-2125		2125-2126		2126-2127		2127-2128		2128-2129		2129-2130		2130-2131		2131-2132		2132-2133		2133-2134		2134-2135		2135-2136		2136-2137		2137-2138		2138-2139		2139-2140		2140-2141		2141-2142		2142-2143		2143-2144		2144-2145		2145-2146		2146-2147		2147-2148		2148-2149		2149-2150		2150-2151		2151-2152		2152-2153		2153-2154		2154-2155		2155-2156		2156-2157		2157-2158		2158-2159		2159-2160		2160-2161		2161-2162		2162-2163		2163-2164		2164-2165		2165-2166		2166-2167		2167-2168		2168-2169		2169-2170		2170-2171		2171-2172		2172-2173		2173-2174		2174-2175		2175-2176		2176-2177		2177-2178		2178-2179		2179-2180		2180-2181		2181-2182		2182-2183		2183-2184		2184-2185		2185-2186		2186-2187		2187-2188		2188-2189		2189-2190		2190-2191		2191-2192		2192-2193		2193-2194		2194-2195		2195-2196		2196-2197		2197-2198		2198-2199		2199-2200		2200-2201		2201-2202		2202-2203		2203-2204		2204-2205		2205-2206		2206-2207		2207-2208		2208-2209		2209-2210		2210-2211		2211-2212		2212-2213		2213-2214		2214-2215		2215-2216		2	
-----------	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	---	--

TGACGGAATCCC GGGCTGGGTATCCTGGTTTNGACAAGATAAACCCCCAGCAANAAATTGGG
GAGCAGGGCAAAACAGTNACGGGCAGCCACATCAAGAAGTTCACCTTNGTTTGNATGGNTC
TGTCAACTCACGCTNTGTTTCGTGATGTTTTGGACACCCAAAGTGTTTGAGAAAATTTTGAT
AGACATNATCGGAGTGGANNTTGCCTTTGCAGAANTTTGNGNTGTTCTTTGCGGATTTTCT
CCTTTTTCCAGTTCAGTCACAGNGAGGGCGCATCTCACCGGGNGGNTGATGACANTGAAG
AAAACCTTTGTCCTTGCCCCAGCTNTTTGGTGCGGATCATTGTCCTNATNGCCAGCCTTGT
GGTCTACCTACCTGGGGGTGCACGGTGCGACCTGGGCGTGGGTTCCCTCCTGGCGGGCA

FIGURE 7

TATTCCCAGTTCCGGTCACGGGGAGGGCGCATNTCACCGGGTGGCTGANGACACTGAAGAAA
ACCTTNGTCCTTGCCCCCAGNTTTGTGNTGCGGATNATCGTCCTCATCGCCAGCCTNGTGGT
CCTACCCTACCTGGGGGTGCACGGTGAGAC

FIGURE 8

GCCCCGCGCCCGGCGCCGGGCGCCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGA
GCCTGCTCCCTGCTCAGCTGCGCGTCTGCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAG
CTGCTGCCCCGCCAGCCGCAACTCCACCGTGAGCCGCTCATCTTCACGTTCTTCCTCTTCC
TGGGGGTGCTGGTGTCCATCATTATGCTGAGCCCCGGGCGTGGAGAGTCAGCTCTACAAGCTG
CCCTGGGTGTGTGAGGAGGGGGCCGGGATCCCCACCGTCTGCAGGGCCACATCGACTGTGG
CTCCCTGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCT
TCTTTTTTACCCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAG
AATGGGTTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACAT
CCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCTGTTGGGCTCCTTCTTCTTCA
TCCTCATCCAGCTGGTGTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGC
AAGGCCGAGGAGTGCGATTCCCGTGCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTCTT
CTACTTGCTGTGATCGCGGCCGTGGCGCTGATGTTTATGTACTACACTGAGCCCAGCGGCT
GCCACGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCT
GCTGTCTGCCCCAAGGTCCAGGACGCCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTCAT
CACCTCTACACCATGTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCA
ACCCCCATTTGCCAACCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAG
ACCCAGTGGTGGGATGCCCCGAGCATTGTGGGCCTCATCATCTTCCTCCTGTGCACCCTCTT
CATCAGTCTGCGCTCCTCAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCC
CACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGTTGGCAGCCTGTGAGGGCCGGGCC
TTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGCT
GGCCTCACTGCACGTATGATGACGCTACCAACTGGTACAAGCCCGGTGAGACCCGGAAGA
TGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTC
TACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTGAGGCAGCCT
CACAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCT
GCCCCCTCCCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCCCAGCTCCAGGACCTG
CCCCTGAGCCGGGCCTTCTAGTCGTAGTGCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCA
GAGCCCCATCCCCCGCCACACCCACACGGTGGAGCTGCCTCTTCCTTCCCCCTCCTCCTGT
TGCCCATACTCAGCATCTCGGATGAAAGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGG
CTGCTGGAGAGAGCGGGGAACCTCCACACAGTGGGGCATCCGGCACTGAAGCCCTGGTGT
CCTGGTCACGTCCCCCAGGGGACCTGCCCCCTTCTGGACTTCGTGCCTTACTGAGTCTCT
AAGACTTTTTTCTAATAACAAGCCAGTGCGTGTAACAAAAA

FIGURE 9

MGACLGACSLLSASCCLCGSAPCILSCCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVE
SOLYKLPWVCEEAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRD
PRAAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFLILIQLVLLIDFAHSW
NQRWLKAEECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFC
VCVSIAAVLPKVQDAQPNSSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVA
GPEGYETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVA
ACEGRAFDNEQDGVITYSYFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVWKICAS
WAGLLLLYLWTLVAPLLLLRNDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 10

GAGCGAGGCCGGGGACTGAAGGTGTGGGTGTGCGAGCCCTCTGGCAGAGGGTTAACCTGGGTC
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCCGCGGCACGTCCGCGAGGACTTGA
AGTCCTGAGCGCTCAAGTTTGTCCGTAGGTGCGAGAGAAGGCCATGGAGGTGCCGCCACCGGC
ACCGCGGAGCTTTCTCTGTAGAGCATTGTGCCTATTTCCCCGAGTCTTTGCTGCCGAAGCTG
TGACTGCCGATTTCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCC
TATTACCCGGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAG
AATTTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGG
TGTATGGGGGAATACCAGCTTTTATTTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCA
GAAATTTATCATAACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTT
CATTCGTTATGGCTGGCGCTGGGGTTGGAGAACTGCAGTGTTTGTGACTATATTCAACACAG
TGAACACTAGTCTGAATGTATACCGAAATAAAGATGCCTTAAGCCATTTTGTAAATTGCAGGA
GCTGTCACGGGAAGTCTTTTTAGGATAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCAT
AATTGGAGCCTTGCTGGGCACCTCTGTAGGAGGCCTGCTGATGGCATTTTCAGAAGTACGCTG
GTGAGACTGTTTCAGGAAAGAAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAACTGGAA
GAGTGGAAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTTACG
GGAAGATGAACCTGAGAATGATGCTAAGAAAATTGAAGCACTGCTAAACCTTCCTAGAAACC
CTTCAGTAATAGATAAAACAAGACAAGGACTTGAAAGTGCTCTGAACTTGAACTCACTGGAGA
GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCTGC
TGACAAATTTAAGTGCTGGTACCTGTGGTGGCAGTGGCTTGCTCTTGCTCTTTTTCTTTCTT
TTTAACTAAGAATGGGGCTGTTGTACTCTCACTTTACTTATCCTTAAATTTAAATACATACT
TATGTTTGTATTAATCTATCAATATATGCATACATGGATATATCCACCCACCTAGATTTTAA
GCAGTAAATAAAACATTTTCGCAAAAGATTAAAGTTGAATTTTACAGTTT

FIGURE 11

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318

><subunit 1 of 1, 285 aa, 1 stop

><MW: 32190, pI: 9.03, NX(S/T): 2

MEVPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRLRELF
GKDEQQRISKDLANICKTAATAGIIGWVYGGIPAFIHAKQQYIEQSQAETIYHNRFDVQSAH
RAATRGFIRYGWRWGWRTAVFVTIFNTVNTSLNVYRNKDALSHFVIAGAVTGSLFRINVGLR
GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

FIGURE 12

CGGAAGTCCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATCTGGATGGGACCGCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGGA
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGAA
TACCAGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATCAT
AACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTCGTTTCATG
GCTGGCGCCGAACC

FIGURE 13

TCAAGTTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGCACCGCGGAGCTT
TTTTCTGTAGAGCATTGTGCCTATTTCCCCGAGTTTTTGCTGCCGAAGCTGTGACTGCCGAT
TCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATTTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGA
ATACCAGCTTTTATTTCATGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATNA
TAACC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Parameter	Estimate	Standard Error	t-Statistic	p-Value
Intercept	0.0000	0.0000	0.0000	0.0000
Age	0.0000	0.0000	0.0000	0.0000
Age squared	0.0000	0.0000	0.0000	0.0000
Age cubed	0.0000	0.0000	0.0000	0.0000
Age quartic	0.0000	0.0000	0.0000	0.0000
Age quintic	0.0000	0.0000	0.0000	0.0000
Age sextic	0.0000	0.0000	0.0000	0.0000
Age septic	0.0000	0.0000	0.0000	0.0000
Age octic	0.0000	0.0000	0.0000	0.0000
Age nonic	0.0000	0.0000	0.0000	0.0000
Age decic	0.0000	0.0000	0.0000	0.0000
Age undecic	0.0000	0.0000	0.0000	0.0000
Age duodecic	0.0000	0.0000	0.0000	0.0000
Age tredecic	0.0000	0.0000	0.0000	0.0000
Age quattuordecic	0.0000	0.0000	0.0000	0.0000
Age quindecic	0.0000	0.0000	0.0000	0.0000
Age sexdecic	0.0000	0.0000	0.0000	0.0000
Age septendecic	0.0000	0.0000	0.0000	0.0000
Age octodecic	0.0000	0.0000	0.0000	0.0000
Age novemdecic	0.0000	0.0000	0.0000	0.0000
Age vigintic	0.0000	0.0000	0.0000	0.0000
Age unguic	0.0000	0.0000	0.0000	0.0000
Age duodevigintic	0.0000	0.0000	0.0000	0.0000
Age tredecimvigintic	0.0000	0.0000	0.0000	0.0000
Age quattuordecimvigintic	0.0000	0.0000	0.0000	0.0000
Age quindecimvigintic	0.0000	0.0000	0.0000	0.0000
Age sexdecimvigintic	0.0000	0.0000	0.0000	0.0000
Age septendecimvigintic	0.0000	0.0000	0.0000	0.0000
Age octodecimvigintic	0.0000	0.0000	0.0000	0.0000
Age novemdecimvigintic	0.0000	0.0000	0.0000	0.0000
Age vigintivigintic	0.0000	0.0000	0.0000	0.0000
Age unguicvigintic	0.0000	0.0000	0.0000	0.0000
Age duodevigintivigintic	0.0000	0.0000	0.0000	0.0000
Age tredecimvigintivigintic	0.0000	0.0000	0.0000	0.0000
Age quattuordecimvigintivigintic	0.0000	0.0000	0.0000	0.0000
Age quindecimvigintivigintic	0.0000	0.0000	0.0000	0.0000
Age sexdecimvigintivigintic	0.0000	0.0000	0.0000	0.0000
Age septendecimvigintivigintic	0.0000	0.0000	0.0000	0.0000
Age octodecimvigintivigintic	0.0000	0.0000	0.0000	0.0000
Age novemdecimvigintivigintic	0.0000	0.0000	0.0000	0.0000
Age vigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age unguicvigintivigintic	0.0000	0.0000	0.0000	0.0000
Age duodevigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age tredecimvigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age quattuordecimvigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age quindecimvigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age sexdecimvigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age septendecimvigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age octodecimvigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age novemdecimvigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age vigintivigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age unguicvigintivigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age duodevigintivigintivigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age tredecimvigintivigintivigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age quattuordecimvigintivigintivigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age quindecimvigintivigintivigintivigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age sexdecimvigintivigintivigintivigintivigintivigintic	0.0000	0.0000	0.0000	0.0000
Age septendecimvigintivigintivigintivigintivigintivigintic	0.0000	0.0000	0.0	

GAGCCGCGCCGCGCGCGCGCGCACCTGCAGCCCCAGGCCCGCCCGCCACCCACCGTCT
GCGTTGCTGCCCCGCTGGGCCAGGCCCAAGGCAAGGACAAAGCAGCTGTCAGGGAACCT
CCGCCGGAGTCGAATTTACGTGCAGCTGCCGGCAACCACAGGTTCCAAGATGGTTTTGCGGGG
GCTTCGCGTGTTCCAAGAACTGCCTGTGCGCCCTCAACCTGCTTTACACCTTGGTTAGTCTG
CTGCTAATTGGAATTGCTGCGTGGGGCATTGGCTTCGGGCTGATTTCCAGTCTCCGAGTGGT
CGGCGTGGTCATTGCAGTGGGCATCTTCTTGTTCTTGATTGCTTTAGTGGGTCTGATTGGAG
CTGTAAACATCATCAGGTGTTGCTATTTTTTTTATATGATTATTCTGTTACTTGTATTTATT
GTTCAGTTTTCTGTATCTTGCGCTTGTTTAGCCCTGAACCAGGAGCAACAGGGTCAGCTTCT
GGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAATCTAAACTGCT
GTGGGTTCCGAAGTGTTAACCCAAATGACACCTGTCTGGCTAGCTGTGTTAAAAGTGACCAC
TCGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGAGATTTGTTGG
TGGCATTGGCCTGTTCTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACCTACAGATACA
GGAACCAGAAAGACCCCCGCGCGAATCCTAGTGCATTCTTTTGATGAGAAAACAAGGAAGAT
TTCCTTTCGTATTATGATCTTGTTCACTTTCTGTAATTTTCTGTTAAGCTCCATTTGCCAGT
TTAAGGAAGGAAACACTATCTGGAAAAGTACCTTATTGATAGTGGAATTATATATTTTTACT
CTATGTTTCTCTACATGTTTTTTTTCTTTCCGTTGCTGAAAAATATTTGAAACTTGTGGTCTC
TGAAGCTCGGTGGCACCTGGAATTTACTGTATTCATTGTGCGGCACTGTCCACTGTGGCCTT
TCTTAGCATTTTTTACCTGCAGAAAACTTTGTATGGTACCACTGTGTTGGTTATATGGTGAA
TCTGAACGTACATCTCACTGGTATAATTATATGTAGCACTGTGCTGTGTAGATAGTTCCTAC
TGGA AAAAGAGTGGAATTTTATTA AAATCAGAAAGTATGAGATCCTGTTATGTTAAGGGAAA
TCCAAATTCCCAATTTTTTTTTGGTCTTTTTTAGGAAAGATTGTTGTGGTAAAAAGTGTTAGTA
TAAAAATGATAATTTACTTG TAGTCTTTTATGATTACACCAATGTATTCTAGAAATAGTTAT
GTCTTAGGAAATTGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTGGTTT
CATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCATCAGAATGGAACGAGTTT
TGAGTAATCAGGAAGTATATCTATATGATCTTGATATTGTTTTATAATAATTTGAAGTCTAA
AAGACTGCATTTTTTAAACAAGTTAGTATTAATGCGTTGGCCACGTAGCAAAAAGATATTTG
ATTATCTTAAAAATTGTTAAATACCGTTTTCATGAAATTTCTCAGTATTGTAACAGCAACTT
GTCAAACCTAAGCATATTTGAATATGATCTCCCATAAATTTGAAATTGAAATCGTATTGTGTG
GCTCTGTATATTCTGTTTAAAAAATTAAAGGACAGAAACCTTCTTTGTGTATGCATGTTTGA
ATTAAAAAGAAAGTAATGGAAG

FIGURE 15

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979

><subunit 1 of 1, 204 aa, 1 stop

><MW: 22147, pI: 8.37, NX(S/T): 3

MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVVIAVGIFLFLIALV
GLIGAVKHHQVLLFFYMIILLLVFIVQFSVSCACLALNQEQQGQLLEVGVWNNTASARNDIQR
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRFVGGIGLFFSFTEILGVWL
TYRYRNQKDPRANPSAFL

Signal Peptide:

amino acids 1-34

Transmembrane domains:

amino acids 47-63, 72-95 and 162-182

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGTNATTTTTTTATATGATTATTCTGTAANT
TGTATTTATTGTTTCAGTTTTNTGTATCTTGCGCTTGTTTAGCCNTGAACCAGGAGCAACAGG
GTCAGNTTNTGGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT
NTAAACTGCTGTGGGTTCGGAAGTGTTAACCCTGACACCTGTNTGGCTAGCTGTGTTAA
AAGTGACCACTNGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGA
GATTTGTTGGTGGCATTGGCCTGTTNTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACC
TACAGATACAGGAACCAG

Variable	Mean	Standard deviation	Minimum	Maximum
Age	34.5	10.5	20	55
Gender	Male	Female	Male	Female
Marital status	Married	Single	Married	Single
Education	High school	College	High school	College
Income	\$10,000	\$20,000	\$10,000	\$20,000
Health status	Good	Fair	Good	Fair
Exercise frequency	Weekly	Monthly	Weekly	Monthly
Stress level	Low	High	Low	High
Smoking status	Smoker	Non-smoker	Smoker	Non-smoker
Alcohol consumption	Regular	Occasional	Regular	Occasional
Family size	2	3	2	3
Home ownership	Owner	Renter	Owner	Renter
Commute time	30 min	45 min	30 min	45 min
Job satisfaction	Satisfied	Dissatisfied	Satisfied	Dissatisfied
Work-life balance	Good	Poor	Good	Poor
Overall well-being	High	Low	High	Low

AATCCCAAATCCCCAATTTTTTTTGGNCTTTTTTAGGGAAAGATGTGTTGTGGTAAAAAGTGT
TAGTATAAAAATGATAATTTACTTGTAGTCTTTTTATGATTACACCAATGTATTCTAGAATAG
TTATGTCTTAGGAAATTGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTG
GTTTCATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCCATCAGAATGGAACG
AGTTTTGAGTAATCCAGGAAGTATATCTATATGATCTTGATATTGTTTTATATAATTTGAAG
TCTAAAAGACTGCATTTTTTAAACAAGTTAGTATTAATGCGTTGGCCACGTAGCAAAAAGAT
ATTTGATTATCTTAAAAATTGTTAAATACCGTTTTTCATGAAAGTTCTCAGTATTGTAACAGC
AACTTGTCAAACCTAAGCATATTTGAATATGATCTCCATAATTTGAAATTGAAATCGTATT
GTGTGGAGGAAATGGCAATCTTATGTGTGCTGAAGGACACAGTAAGAGCACCAAGTTGTGCC
CCACTTGC

FIGURE 18

ATGATTATTCTGTTACTTGTATTTATTG TTCAGTTTTATGGTATCTTGCGCTTGTTTAGCCC
CTGAAACCAGGAGCAACAGGGNNCAGCTTCCTGGAGGTTGGTTGGCAACAATCACGGCCAAG
TGACTCCGCAAATGACATCCCAGAGAAATCCTAAACTGCTGTGGGTTCGGAAGTGTTAACCC
AAATGACACCTGTCTGGCTNGCTGTGTTAAAAGTGACCACTCGTGCTCGCCATGTGCTCCAA
TCATAGGAGAATATGC

FIGURE 19

CAGTCACCATGAAGCTGGGCTGTGTCCTCATGGCCTGGGCCCTCTACCTTTCCCTTGGTGTG
CTCTGGGTGGCCAGATGCTACTGGCTGCCAGTTTGTGAGACGCTGCAGTGTGAGGGACCTGT
CTGCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTTGACTGATGCAAGGGAAGCTGGCT
TCCAGGTCAAGGCCTACACTTTTCAGTGAACCCTTCCACCTGATTGTGTCTTATGACTGGCTG
ATCCTCCAAGGTCCAGCCAAGCCAGTTTTTGAAGGGGACCTGCTGGTTCTGCGCTGCCAGGC
CTGGCAAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCC
CCGGGCCTAACAGGGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCAC
TGCAGTGGCATCTTCCAGAGCCCTGGTCTGGGATCCCAGAAACAGCATCTGTTGTGGCTAT
CACAGTCCAAGAACTGTTTCCAGCGCCAATTCTCAGAGCTGTACCCTCAGCTGAACCCCAAG
CAGGAAGCCCCATGACCCTGAGTTGTGAGACAAAGTTGCCCTGCAGAGGTGAGCTGCCCGC
CTCCTCTTCTCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGGCTCTCCTCAGAATT
CCAGATCCCCACAGCTTCAGAAGATCACTCCGGGTCTACTGGTGTGAGGCAGCCACTGAGG
ACAACCAAGTTTGGAAACAGAGCCCCCAGCTAGAGATCAGAGTGCAGGGTGCTTCCAGCTCT
GCTGCACCTCCCACATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAACCTGCTCCTGA
GGAGGCCCTGGGCCTCTGCCTCCGCCGCCAACCCCATCTTCTGAGGATCCAGGCTTTTCTT
CTCCTCTGGGGATGCCAGATCCTCATCTGTATCACCAGATGGGCCTTCTTCTCAAACACATG
CAGGATGTGAGAGTCTCCTCGGTACCTGCTCATGGAGTTGAGGGAATTATCTGGCCACCA
GAAGCCTGGGACCACAAAGGCTACTGCTGAATAGAAAGTAAACAGTTCATCCATGATCTCACT
TAACCACCCCAATAAATCTGATTCTTTATTTTCTCTTCTGCTGCTGCACATATGCATAAGTA
CTTTTACAAGTTGTCCCAGTGTTTTGTTAGATAATGTAGTTAGGTGAGTGTAAATAAATTT
ATATAAAGTGAGAATTAGAGTTTAGCTATAATTGTGTATTCTCTCTTAACACAACAGAATTC
TGCTGTCTAGATCAGGAATTTCTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAA
CTAATGGAAGTGGATTGAATACAGCAGTCTCAACTGGGGGCAATTTTGCCCCCAGAGGACA
TTGGGCAATGTTTGGAGACATTTTGGTCATTATACTTGGGGGGTGGGGGATGGTGGGATGT
GTGTCTACTGGCATCCAGTAAATAGAAGCCAGGGGTGCCGCTAAACATCCTATAATGCACAG
GGCAGTACCCACAAACGAAAAATAATCTGGCCCAAATGTGAGTTGTACTGAGTTTGAGAAA
CCCCAGCCTAATGAAACCCTAGGTGTTGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTAT
TATCTCTTTCCAGCCTCATTGAGCTATTCTTACTGACATACCAGTCTTTAGCTGGTGCTATG
GTCTGTTCTTTAGTTCTAGTTTGTATCCCCTCAAAGCCATTATGTTGAAATCCTAATCCCC
AAGGTGATGGCATTAAAGAAGTGGGCCTTTGGGAAGTGATTAGATCAGGAGTGCAGAGCCCTC
ATGATTAGGATTAGTGCCCTTATTTAAAAAGGCCCCAGAGAGCTAACTCACCTTCCACCAT
ATGAGGACGTGGCAAGAAGATGACATGTATGAGAACCACAAAAACAGCTGTCGCCAAACACCG
ACTCTGTCGTTGCCTTGATCTTGAACCTCCAGCCTCCAGAACTATGAGAAATAAAATTCTGG
TTGTTTGTAGCCTAA

FIGURE 20

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594

><subunit 1 of 1, 359 aa, 1 stop

><MW: 38899, pI: 5.21, NX(S/T): 0

MKLGCVLMAWALYLSLGVLWVAQMLLAASFETLQCEGPVCTEESSCHTEDDLTDAREAGFQV
KAYTFSEPFHLIVSYDWLILQGPAKPVFEGDLLVLRQAWQDWPLTQVTFYRDGSALGPPGP
NREFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVVAITVQELFPAPILRAVPSAEPQAGS
PMTLSCQTKLPLQRSAARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHSGSYWCEAATEDNQ
VWKQSPQLEIRVQGASSAAPPPTLNPAPQKSAAPGTAPEEAPGPLPPPPTPSSSEDPGFSSPL
GMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE

Signal sequence:

amino acids 1-17

Leucine zipper pattern sequence:

amino acids 12-33

Protein kinase C phosphorylation site:

amino acids 353-355

Figure 20: 353-355

FIGURE 21

CCCACGCGTCCGCCCACGCGTCCGCCCACGGGTCCGCCCACGCGTCCGGGCCACCAGAAGTT
TGAGCCTCTTTGGTAGCAGGAGGCTGGAAGAAAGGACAGAAGTAGCTCTGGCTGTGATGGGG
ATCTTACTGGGCCTGCTACTCCTGGGGCACCTAACAGTGGACACTTATGGCCGTCCCATCCT
GGAAGTGCCAGAGAGTGTAAACAGGACCTTGGAAAGGGGATGTGAATCTTCCCTGCACCTATG
ACCCCTGCAAGGCTACACCCAAGTCTTGGTGAAGTGGCTGGTACAACGTGGCTCAGACCCT
GTCACCATCTTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGCAAAGTACCAGGGCCG
CCTGCATGTGAGCCACAAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCTTGGAGATGG
ATGACCGGAGCCACTACACGTGTGAAGTCACCTGGCAGACTCCTGATGGCAACCAAGTCGTG
AGAGATAAGATTACTGAGCTCCGTGTCCAGAACTCTCTGTCTCCAAGCCCACAGTGACAAC
TGGCAGCGGTTATGGCTTCACGGTGCCCCAGGGAATGAGGATTAGCCTTCAATGCCAGGCTC
GGGGTTCTCCTCCCATCAGTTATATTTGGTATAAGCAACAGACTAATAACCAGGAACCCATC
AAAGTAGCAACCCTAAGTACCTTACTCTTCAAGCCTGCGGTGATAGCCGACTCAGGCTCCTA
TTTCTGCACTGCCAAGGGCCAGGTTGGCTCTGAGCAGCACAGCGACATTGTGAAGTTTGTGG
TCAAAGACTCCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATACCCC
TTGAAAGCAACATCTACAGTGAAGCAGTCCTGGGACTGGACCACTGACATGGATGGCTACCT
TGGAGAGACCAGTGCTGGGCCAGGAAAGAGCCTGCCTGTCTTTGCCATCATCCTCATCATCT
CCTTGTGCTGTATGGTGGTTTTTACCATGGCCTATATCATGCTCTGTTCGGAAGACATCCCAA
CAAGAGCATGTCTACGAAGCAGCCAGGTAAGAAAGTCTCTCCTCTTCCATTTTTTGACCCCGT
CCCTGCCCTCAATTTTGATTACTGGCAGGAAATGTGGAGGAAGGGGGGTGTGGCACAGACCC
AATCCTAAGGCCGGAGGCCTTCAGGGTCAGGACATAGCTGCCTTCCCTCTCTCAGGCACCTT
CTGAGGTTGTTTTGGCCCTCTGAACACAAAGGATAATTTAGATCCATCTGCCTTCTGCTTCC
AGAATCCCTGGGTGGTAGGATCCTGATAATTAATTGGCAAGAATTGAGGCAGAAGGGTGGGA
AACCAGGACCACAGCCCCAAGTCCCTTCTTATGGGTGGTGGGCTCTTGGGCCATAGGGCACA
TGCCAGAGAGGCCAACGACTCTGGAGAAACCATGAGGGTGGCCATCTTCGCAAGTGGCTGCT
CCAGTGATGAGCCAACCTTCCCAGAATCTGGGCAACAACCTACTCTGATGAGCCCTGCATAGGA
CAGGAGTACCAGATCATCGCCCAGATCAATGGCAACTACGCCCGCCTGCTGGACACAGTTCC
TCTGGATTATGAGTTTCTGGCCACTGAGGGCAAAGTGTCTGTAAATAATGCCCCATTAGGC
CAGGATCTGCTGACATAATTGCCTAGTCAGTCCTTGCCTTCTGCATGGCCTTCTTCCCTGCT
ACCTCTCTTCCCTGGATAGCCCCAAGTGTCCGCCTACCAACACTGGAGCCGCTGGGAGTCACT
GGCTTTGCCCTGGAATTTGCCAGATGCATCTCAAGTAAGCCAGCTGCTGGATTTGGCTCTGG
GCCCTTCTAGTATCTCTGCCGGGGGCTTCTGGTACTCCTCTCTAAATACCAGAGGGAAGATG
CCCATAGCACTAGGACTTGGTCATCATGCCTACAGACACTATTCAACTTTGGCATCTTGCCA
CCAGAAGACCCGAGGGAGGCTCAGCTCTGCCAGCTCAGAGGACCAGCTATATCCAGGATCAT
TTCTCTTTCTTCAAGGCCAGACAGCTTTTAATTGAAATTGTTATTTACAGGCCAGGGTTCA
GTTCTGCTCCTCCACTATAAGTCTAATGTTCTGACTCTCTCCTGGTGCTCAATAAATATCTA
ATCATAACAGC

[illegible]

><MW: 35544, pI: 8.51, NX(S/T): 0

amino acids 282-300

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLVSYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDLTLCESSTGTEPIVYVWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMSSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSSGSRSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

FIGURE 25

GTCGTTCCCTTTGCTCTCTCGCGCCAGTCCTCCTCCCTGGTTCTCCTCAGCCGCTGTGCGAGGAGAGCACC CGGA
 GACGCGGGCTGCAGTCGCGGCGGGCTTCTCCCCGCTGGGCGGCCTCGCCGCTGGGCAGGTGCTGAGCGCCCTAG
 AGCCTCCCTTGCCGCTCCTCCTCTGCCCGCCGAGCAGTGACATGGGGTGTGGAGGTAGATGGGCTCCCG
 GCCCCGGAGGCGCGGTGGATGCGGCGCTGGGCAGAAGCAGCCGCGATTCCAGCTGCCCCGCGCGCCCCGGGCG
 CCCCTGCGAGTCCCCGTTTACGCCATGGGGACCTCTCCGAGCAGCAGCACC GCCCTCGCCTCCTGCAGCCGCATC
 GCCCGCCGAGCCACAGCCACGATGATCGCGGGCTCCTTTCTCCTGCTTGGATTCTTTAGCACCACCACAGCTCAG
 CCAGAACAGAAGGCTCGAATCTCATTTGGCACATACCGCCATGTTGACCGTGCCACCGGCCAGGTGCTAACCTGT
 GACAAGTTCAGCAGGAACCTATGTCTCTGAGCATTGTACCAACACAAGCCTGCGCGTCTGCAGCAGTTGCCCT
 GTGGGGACCTTTTACCAGGCATGAGAATGCGATAGAGAAATGCCATGACTGTAGTCAGCCATGCCCATGGCCAAATG
 ATTGAGAAATTACCTTTGTGCTGCCTTGACTGACCGAGAATGCACTTGCCACCTGGCATGTTCCAGTCTAACGCT
 ACCTGTGCCCCCATAACGTTGTCTGTGGGTTGGGGTGTGCGGAAGAAAGGGACAGAGACTGAGGATGTGCGG
 TGTAAAGCAGTGTGCTCGGGGTACCTTCTCAGATGTGCCTTCTAGTGTGATGAAATGCAAAGCATACACAGACTGT
 CTGAGTCAGAACCCTGGTGGTGATCAAGCCGGGGACCAAGGAGACAGACAACGCTCTGTGGCACACTCCCGTCTTTC
 TCCAGCTCCACCTCACCTTCCCTGGCACAGCCATCTTCCACGCCCTGAGCACATGGAAACCCATGAAGTCCCT
 TCCTCCACTTATGTTCCCAAAGGCATGAACCTCAACAGAATCCAACCTCTTCTGCCTCTGTTAGACCAAAGGTACTG
 AGTAGCATCCAGGAAGGGACAGTCCCTGACAACAAGCTCAGCAAGGGGAAGGAAGACGTGAACAAGACCCCTC
 CCAAACCTTCAGGTAGTCAACCACCAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGAG
 GCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGGGACATCCTAGACAGAACCTACACAAG
 CATTTTGACATCAATGAGCATTTGCCCTGGATGATTGTGCTTTTCTGCTGCTGGTGTCTGTGGTGATTGTGGTG
 TGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGGATCCCAGTGCCATTGTGAAAAGGCA
 GGGCTGAAGAAATCCATGACTCCAACCCAGAACC GGAGAAAATGGATCTACTACTGCAATGGCCATGGTATCGAT
 ATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCTATCAGTTTCTTTGCAATGCCAGTGAG
 AGGGAGGTTGTGCTTTCTCCAATGGGTACACAGCCGACCACGAGCGGGCTTACGCAGCTCTGCAGCACTGGACC
 ATCCGGGGCCCCGAGGCCAGCCTCGCCAGCTAATTAGCGCCCTGCGCCAGCACCGGAGAAAACGATGTTGTGGAG
 AAGATTCGTGGGCTGATGGAAGACACCACCCAGCTGGAAACTGACAACTAGCTCTCCCGATGAGCCCCAGCCCCG
 CTTAGCCCCGAGCCCCATCCCCAGCCCCAACGCGAACTTGAGAAATCCGCTCTCCTGACGGTGGAGCCTTCCCCA
 CAGGACAAGAACCAAGGGCTTCTTCGTGGATGAGTCGGAGCCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTCC
 TCCGCGCTGAGCAGGAACGGTTCTTTTATTACCAAAGAAAAGAAGGACACAGTGTGCGGCAGGTACGCCCTGGAC
 CCCTGTGACTTGCAGCCTATCTTTGATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGAG
 ATTCCCCAGGCTGAGGACAACTAGACCGCTATTGCAAATTATTGGAGTCAAGAGCCAGGAAGCCAGCCAGACC
 CTCCTGGACTCTGTTTATAGCCATCTTCTGACCTGCTGTAGAACATAGGGATACTGCATTCTGGAAATTACTCA
 ATTTAGTGGCAGGGTGGTTTTTTAATTTTCTTCTGTTTCTGATTTTTGTGTTTGGGGTGTGTGTGTGTTTGT
 GTGTGTGTGTGTGTGTGTGTGTGTGTTTAAACAGAGAATATGGCCAGTGCCTTGAGTTCTTTCTCCTTCTC
 TCTCTCTCTTTTTTTTTTAAATAACTCTTCTGGGAAGTTGGTTTATAAGCCTTTGCCAGGTGTAACGTGTTGTGAA
 ATACCCACCACTAAAGTTTTTTAAGTTCCATATTTTCTCCATTTTGCCTTCTTATGTATTTTCAAGATTATTCTG
 TGCATTTTAAATTTACTTAACTTACCATAAATGCAGTGTGACTTTTCCACACACTGGATTGTGAGGCTCTTAAC
 TTCTTAAAGTATAATGGCATCTTGTGAATCCTATAAGCAGTCTTTATGTCTCTTAACATTACACCTACTTTTT
 AAAAACAAATATTATTACTATTTTATTATTGTTTGTCTTTATAAAATTTCTTAAAGATTAAAGAAAATTTAAGA
 CCCCATTGAGTTACTGTAATGCAATTCAACTTTGAGTTATCTTTTAAATATGCTTTGTATAGTTTCATATTTCATGG
 CTGAAACTTGACCACACTATTGCTGATTGTATGGTTTTTACCTGGACACCGTGTAGAATGCTTGATTACTTGTAC
 TCTTCTTATGCTAATATGCTCTGGGCTGGAGAAATGAAATCCTCAAGCCATCAGGATTTGCTATTTAAGTGGCTT
 GACAACTGGGCCACCAAAGAACTTGAACCTTACCTTTTAGGATTTGAGCTGTTCTGGAACACATTGCTGCACCTT
 GGAAAGTCAAAATCAAGTGCCAGTGGCGCCCTTTCCATAGAGAATTTGCCAGCTTTGCTTTAAAGATGTCTTG
 TTTTTTATATACATAATCAATAGGTCCAATCTGCTCTCAAGGCCTTGGTCTGGTGGGATTCCCTTACCAATT
 ACTTTAATTAAAAATGGCTGCAACTGTAAAGAACCTTGTCTGATATATTTGCAACTATGCTCCCATTTACAAATG
 TACCTTCTAATGCTCAGTTGCCAGGTTCCAATGCAAAGGTGGCGTGGACTCCCTTTGTGTGGGTGGGGTTTTGTGG
 GTAGTGGTGAAGACCGATATCAGAAAAATGCCTTCAAGTGACTAATTTATTATAAACATTAGGTGTTTGTTA
 AAAAAAAA

[illegible]

```
><subunit 1 of 1, 655 aa, 1 stop
```

[illegible]

amino acids 1-41

amino acids 350-370

FIGURE 27

ATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCCGTGGTGCCATCTACATTTTTTGGGA
CTCGGGAATTATGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCAC
CATGGGGGAAAATGATCCGCCTGCTGTTGAAGCCCCCTTCTCATTCCGATCGCTTTTTTGGCC
TTGATGATTTGAAAATAAGTCCTGTTGCACCAGATGCAGATGCTGTTGCTGCACAGATCCTG
TCACTGCTGCCATTGAAGTTTTTTTCCAATCATCGTCATTGGGATCATTGCATTGATATTAGC
ACTGGCCATTGGTCTGGGCATCCACTTCGACTGCTCAGGGAAGTACAGATGTCGCTCATCCT
TTAAGTGTATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAAGACGGGGAGGAC
GAGTACCGCTGTGTCCGGGTGGGTGGTCAGAAATGCCGTGCTCCAGGTGTTTACAGCTGCTTC
GTGGAAGACCATGTGCTCCGATGACTGGAAGGGTCACTACGCAAATGTTGCCTGTGCCAAC
TGGGTTTTCCCAAGCTATGTGAGTTCAGATAACCTCAGAGTGAGCTCGCTGGAGGGGCAGTTC
CGGGAGGAGTTTGTGTCCATCGATCACCTCTTGCCAGATGACAAGGTGACTGCATTACACCA
CTCAGTATATGTGAGGGAGGGATGTGCCCTCTGGCCACGTGGTTACCTTGCAGTGCACAGCCT
GTGGTCATAGAAGGGGCTACAGCTCACGCATCGTGGGTGGAAACATGTCCTTGCTCTCGCAG
TGGCCCTGGCAGGCCAGCCTTCAGTTCAGGGGTACCACCTGTGCGGGGGCTCTGTATCAC
GCCCCCTGTGGATCATCACTGCTGCACACTGTGTTTTATGACTTGTACCTCCCCAAGTCATGGA
CCATCCAGGTGGGTCTAGTTTCCCTGTTGGACAATCCAGCCCCATCCCACTTGGTGGAGAAG
ATTGTCTACCACAGCAAGTACAAGCCAAAGAGGCTGGGCAATGACATCGCCCTTATGAAGCT
GGCCGGGCCACTCACGTTCAATGAAATGATCCAGCCTGTGTGCCTGCCAACTCTGAAGAGA
ACTTCCCCGATGGAAGAGTGTGCTGGACGTGAGGATGGGGGGCCACAGAGGATGGAGGTGAC
GCCTCCCCTGTCTGAACCACGCGGCCGTCCCTTTGATTTCCAACAAGATCTGCAACCACAG
GGACGTGTACGGTGGCATCATCTCCCCCTCCATGCTCTGCGCGGGCTACCTGACGGGTGGCG
TGGACAGCTGCCAGGGGGACAGCGGGGGGCCCTGGTGTGTCAAGAGAGGAGGCTGTGGAAG
TTAGTGGGAGCGACCAGCTTTGGCATCGGCTGCGCAGAGGTGAACAAGCCTGGGGTGTACAC
CCGTGTACCTCCTTCCTGGACTGGATCCACGAGCAGATGGAGAGAGACCTAAAAACCT**TGAA**
GAGGAAGGGGACAAGTAGCCACCTGAGTTCCTGAGGTGATGAAGACAGCCCGATCCTCCCCT
GGACTCCCGTGTAGGAACCTGCACACGAGCAGACACCTTGGAGCTCTGAGTTCGGGCACCA
GTAGCAGGCCCCGAAAGAGGCACCTTCCATCTGATTCCAGCACAACTTCAAGCTGCTTTTT
GTTTTTTGTTTTTTTTGAGGTGGAGTCTCGCTCTGTTGCCAGGCTGGAGTGCAGTGGCGAAA
TCCCTGCTCACTGCAGCCTCCGCTTCCCTGGTTCAAGCGATTCTCTTGCCTCAGCTTCCCCA
GTAGCTGGGACCACAGGTGCCCGCCACCACACCAACTAATTTTTGTATTTTTTAGTAGAGAC
AGGGTTTCACCATGTTGGCCAGGCTGCTCTCAAACCCCTGACCTCAAATGATGTGCCTGCTT
CAGCCTCCCACAGTGCTGGGATTACAGGCATGGGCCACCACGCCTAGCCTCAGCTCCTTTC
TGATCTTCACTAAGAACAAAAGAAGCAGCAACTTGCAAGGGCGGCCTTCCCCTGGTCCAT
CTGGTTTTCTCTCCAGGGTCTTGCAAAATTCTGACGAGATAAGCAGTTATGTGACCTCACG
TGCAAAGCCACCAACAGCCACTCAGAAAAGACGCACCAGCCCAGAAGTGCAGAACTGCAGTC
ACTGCACGTTTTTCATCTCTAGGGACCAGAACCACCAACCCACCTTTCTACTTCCAAGACTTAT
TTTCACATGTGGGGAGGTTAATCTAGGAATGACTCGTTTAAGGCCTATTTTCATGATTCTT
TGTAGCATTTGGTGCTTGACGTATTATTGTCTTTGATTCAAATAATATGTTTCCTTCCCT
CATTGTCTGGCGTGTCTGCGTGGACTGGTGACGTGAATCAAATCATCCACTGAAA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234

><subunit 1 of 1, 453 aa, 1 stop

><MW: 49334, pI: 6.32, NX(S/T): 1

MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVVRVGGQNAVLQVFTAAS
WKTMCSDDWKGHYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSIDHLLPDDKVTALHH
SVYVREGCASGHVVTLQCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT
PLWIIITAAHCVDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYHSKYKPKRLGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAPPLISNKICNHR
DVYGGIISPSMLCAGYLTGGVDSCQGDSSGGLVCQERRLWKLVGATSFQIGCAEVNKPVGYYT
RVTSFLDWIHEQMERDLKT

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

FIGURE 29

CCCACGCGTCCGTCTAGTCCCCGGGCCAACTCGGACAGTTTGCTCATTTATTGCAACGGTCAAGGCTGGCTTGT
GCCAGAACGGCGCGCGCGCGCACGACGACACACGCGGGGAAACTTTTTTAAAAATGAAAGGCTAGAAGA
GCTCAGCGGGCGCGCGGGCGCTGCGCGAGGGCTCCGGAGCTGACTCGCCGAGGCAGGAAATCCCTCCGGTCGCGA
CGCCCGGCCCCGGCTCGGCGCCCGGTGGGATGGTGCAGCGCTCGCCGCGGGCCGAGAGCTGCTGCACCTGAAG
GCCGGCGACGATGGCAGCGCGCCCGCTGCCCGTGTCCCCCGCCGCGCCCTCCTGCTCGCCCTGGCCGGTGTCT
GCTCGCGCCCTGCGAGGCCCGAGGGGTGAGCTTATGGAACCAAGGAAGAGCTGATGAAGTTGTCAAGTGCCTCTGT
TCGGAGTGGGGACCTCTGGATCCCACTGAAGAGCTTCGACTCCAAGAATCATCCAGAAGTGTGAATATTCGACT
ACAACGGGAAAGCAAAGAACTGATCATAAATCTGGAAAGAAATGAAGGTCTCATTGCCAGCAGTTTCACGGAAAC
CCACTATCTGCAAGACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGGTCACTGTTACTACCATGGACATGT
ACGGGGATATTTCTGATTCAGCAGTCACTCTCAGCACGTGTTCTGGTCTCAGGGGACTTATTGTGTTTGAAAATGA
AAGCTATGTCTTAGAACCAATGAAAAGTGCAACCAACAGATACAACTCTTCCAGCGAAGAAGCTGAAAAGCGT
CCGGGGATCATGTGGATCACATCACACACACCAAACTCGCTGCAAGAATGTGTTTCCACCACCCTCTCAGAC
ATGGGCAAGAAGGCATAAAAGAGAGACCCCTCAAGGCAACTAAGTATGTGGAGCTGGTGTATCGTGGCAGACAACCG
AGAGTTTTCAGAGGCAAGGAAAAGATCTGGAAAAAGTTAAGCAGCGATTAAATAGAGATTGCTAATCACGTTGACAA
GTTTTACAGACCACTGAACATTCTGGATCGTGTGGTAGGCGTGGAAGTGTGAATGACATGGACAAAATGCTCTGT
AAGTCAGGACCCATTACACAGCCTCCATGAATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAATCCCA
TGACAATGCGCAGCTTGTCACTGGGGTTTATTTCCAAGGGACCACCATCGGCATGGCCCCAATCATGAGCATGTG
CACGGCAGACCAGTCTGGGGGAATTGTCTAGGACATTACAGACAATCCCTTGGTGCAGCCGTGACCTGGCACA
TGAGCTGGGCCACAATTTCTGGGATGAATCATGACACACTGGACAGGGGCTGTAGCTGTCAAATGGCGGTTGAGAA
AGGAGGCTGCATCATGAACGCTTCCACCGGGTACCCATTTCCCATGGTGTTCAGCAGTTGCAGCAGGAAGGACTT
GGAGACCAGCCTGGAGAAAAGGAATGGGGGTGTGCCTGTTTAACTGCGGGAAGTCAGGGAGTCTTTCGGGGGCCA
GAAGTGTGGGAACAGATTTGTGGAAGAAGGAGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATGAATCGCTG
CTGCAATGCCACCACCTGTACCTGAAGCCGACGCTGTGTGCGCACATGGGCTGTGCTGTGAAGACTGCCAGCT
GAAGCCTGCAGGAACAGCGTGCAGGGACTCCAGCAACTCCTGTGACCTCCAGAGTTCTGCACAGGGGCCAGCCC
TCACTGCCCAGCCAATGTGTACCTGCACGATGGGCACTCATGTCAAGATGTGGACGGCTACTGCTACAATGGCAT
CTGCCAGACTCACGAGCAGCAGTGTGTACGCTCTGGGGACCAGGTGCTAAACCTGCCCTGGGATCTGCTTTGA
GAGAGTCAATTTCTGAGGTGATCCTTATGGCAACTGTGGCAAGTCTCGAAGAGTTCTTTGCCAAATGCGAGAT
GAGAGATGCTAAATGTGGAATAATCCAGTGTCAAGGAGGTGCCAGCCGGCCAGTCAATTGGTACCAATGCCGTTTC
CATAGAAACAAACATCCCTCTGCAGCAAGGAGGCCGATTCTGTGCCGGGGGACCCACGTGTACTTGGGCGATGA
CATGCCGGACCCAGGGCTTGTGCTTGCAGGCACAAAGTGTGCAGATGGAAAAATCTGCCTGAATCGTCAATGTCA
AAATATTAGTGTCTTTGGGGTTCACGAGTGTGCAATGCAGTGCCACGGCAGAGGGGTGTGCAACAACAGGAAGAA
CTGCCAATGCGAGGCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGCTTTGGAGGAAGCACAGACAGCGGCC
CATCCGGCAAGCAGAAGCAAGGCAGGAAGCTGCAGAGTCCAACAGGGAGCGCGGCCAGGGCCAGGAGCCCGTGGG
ATCGCAGGAGCATGCGTCTACTGCCTCACTGACACTCATCTGAGCCCTCCCATGACATGGAGACCGTGACCACTG
CTGCTGCAGAGGAGGTACGCGTCCCCAAGGCCCTCTGTGACTGGCAGCATTGACTCTGTGGCTTTGCCATCGTT
TCCATGACAACAGACACAACACAGTTCTCGGGCTCAGGAGGGGAAGTCCAGCCTACCAGGCACGTCTGCAGAAA
CAGTGCAAGGAAGGGCAGCGACTTCTGGTTGAGCTTCTGCTAAAAATGACATGCTTCAGTGTCTCCTGAG
AGAGTAGCAGGTTACCACTCTGGCAGGCCCCAGCCCTGCAGCAAGGAGGAAGAGGACTCAAAAGTCTGGCCTTTC
ACTGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTTGGGCCCAAGTGTCCCCTTTCCCACTGACACCTCAGCCT
TGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTCTGATATGGCTTTTAGCATTTATTATATGAAAAT
AGCAGGGTTTTAGTTTTTAATTTATCAGAGACCTGCCACCCATTCCATCTCCATCCAAGCAAACCTGAATGGCAA
TGAAACAACTGGAGAAGAAGGTAGGAGAAAGGGCGGTGAACTCTGGCTCTTTGCTGTGGACATGCGTGACCAGC
AGTACTCAGGTTTGGAGGTTTGCAGAAAGCCAGGGAACCCACAGAGTCACCAACCTTCAATTAAACAAGTAAGAA
TGTTAAAAAGTGAAAAAATGTAAGAGCCTAACTCCATCCCCCGTGGCCATTACTGCATAAAATAGAGTGCATTT
GAAAT

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624

><subunit 1 of 1, 735 aa, 1 stop

><MW: 80177, pI: 7.08, NX(S/T): 5

MAARPLPVSPARALLLAGALLAPCEARGVSLWNQGRADEVVSASVRSGLWIPVKSFDSK
NHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGHCCYYHGHVRG
YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHHNTPNLA
AKNVFPPPSQTWARRHKRETLKATKYVELVIVADNREFQRQGDLEKVKQRLIEIANHVDKF
YRPLNIRIVLVGVEVWMDMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQLVSGVYFQG
TTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFGMNHDTLDRGCSCQMAVEK
GGCIMNASTGYPFPMVFSSCSRKDLETSLEKGMGVCLFNLPEVRESFQQKCGNRFVEEGEE
CDCGEPEECMNRCNATTCTLKPDVCAHGLCCEDCQLKPAGTACRDSSNSCDLPEFCTGAS
PHCPANVYLHDGHSCQDVDGYCYNGICQTHEQQCVTLWGPAGKAPPGICFERVNSAGDPYGN
CGKVSKSSFAKCEMRDAKCGKIQCGGASRPVIGTNAVSIETNIPLQQGGRILCRGTHVYLG
DDMPDPGLVLAGTKCADGKICLNRQCQNISVFGVHECAMQCHGRGVCNNRKNCHCEAHWAPP
FCDKFGFGGSTDSGPQRQAEARQEAESNRERGGQEPVGSQEHASTASLTLI

Signal peptide:

amino acids 1-28

FIGURE 31

TCCCAAGGCTTCTTGGATGGCAGATGATTNTGGGGTTTTGCATTGTTTCCCTGACAACGAAA
ACAAAACAGTTTTGGGGGTTTCAGGAGGGGAANTCCAGCCTACCCAGGAAGTTTGCAGAAACA
GTGCAAGGAAGGGCAGGANTTCCTGGTTGAGNTTTTTGNTAAAACATGGACATGNTTCAGTG
CTGCTCNTGAGAGAGTAGCAGGTTACCACTTTTGGCAGGCCCCAGCCCTGCAGCAAGGAGGA
AGAGGACTCAAAAGTTTGGCCTTTCACTGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTT
GGGCCCAGTGTCCTTTTCCCCAGTGACACCTCAGCCTTGGCAGCCCTGATAACTGGTNTNT
GGCTGCAANTTAATGCTNTGATATGGCTTTTAGCATTTATTATATGAAAATAGCAGGGTTTT
AGTTTTTAATTTATCAGAGACCCTGCCACCCATTCCATNTCCATCCAAG

FIGURE 32

CATCCTGCAACATGGTGAAACCACGCCTGGCTAATTTTGTGTATTTTGGTAGAGATGGGA
TTTCACCGTGTTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCCGCTCGGCCTCCC
AAAGTGCTGGGATTACAGGCGAGTGCAACCACACCCGGCCACAACTTTTTAAGAAGTTAAT
GAAACCATAACCTTTTACATTTTAAATGACAGGAAAATGCTCACAATAATTGTTAACCCAAAA
TTCTGGATACAAAAGTACAATCTTTACTGTGTAAATACATGTATATGTACTATATGAAAATA
TACCAAATATCAATAATACTTATCTCTGGGTAAAAACCTCTTCTCATACCTGTGCTAACAA
CTTTTAACAAAAAATTTGCATCACTTTTAAAGAATCAAGAAAAATTTCTGAAGGTCATATGGG
ACAGAAAAAAAACCAAGGGAAAAATCACGCCACTTGGGAAAAAAGATTGGAATCTGCCT
TTTTATAGATTTGTAATTAATAAGGTCCAGGCTTCTAAGCAACTTAAATGTTTTGTTTCGA
AACAAAGTACTTGTCTGGATGTAGGAGGAAAGGGAGTGATGTCACTGCCATTATGATGCCCC
TTGAATATAAGACCCTACTTGCTATCTCCCCTGACCAGCCAGGAGCCACCCATCCTCCAGC
ACACTGAGCAGCAAGCTGGACACACGGCACACTGATCCAAATGGGTAAGGGGATGGTGGCGA
TGCTCATTCTGGGTCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTTCTTCATTGT
CCTTTAACCAGTATGCCGGAAGCTACTGCAGCCGAAACCACAAAGCCCTCCAACAGTGCCCT
ACAGCCTACAGCCGGTCTCCTTGTTGGTCTTGCTTGCCCTTCTACATCTCTACCATTAAAGAGG
CAGGTCAAGAAACAGCTACAGTTCTCCAACCCATACACTAAAACCGAATCCAAATGGTGCCT
AGAAGTTCAATGTGGCAAGGAAAAAAACCAGGTCTTCATCAAATCTACTAATTTCACTCCTT
ATTAACAGAGAAACGCTTGAGAGTCTCAAACCTGGACTGGTTTAAAGAGCATCTGAAGGATTT
GACTAGATGATAAATGCCTGTACTCCAGTACTTTGGGAGGCCTAGGCCGGCGGATCACCTG
AGGTCAGGAGTTTGAGACTAACCTGGCCAAAATGGTGAAACCCCATCTGTACTAAAAATACA
AATATTGACTGGGCGTGGTGGTGAGTGCCTGTGATCCCAGCTACTCAGGTGGCTGAAGCAGG
ACAATCACTTGAACCTCAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGCTACTGCACTCTA
GCCTAGCCTGGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAAAAGCCAAGTGCAGTGGCT
CACGCCTGTAATCCCGGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAGGTCAGGAGATCA
AGACCATCCTGGCTAATACAGTGAAACCCTGTCTCTACTAAAAATACAAAAAATTAGCCGGG
GATGGTGGCAGGCACCTGGAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATAGCGTGAA
CTCAGGAGGCGGAGCTTGCAGTGAGCCGAGATTGCGCTACTGCACTCCAGCCTGGGCGACAG
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAA

[illegible]

amino acids 15-27

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2423	2424	2425	2426	2427	2428	2429	2430	2431	2432	2433	2434	2435	2436	2437	2438	2439	2440	2441	2442	2
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

CGCCGCGCGAGAGCGCGCCACGCCCGCCGCGCATGCCCGCGCGCCAGGACGCCTTCTCCGCTGCTGGCCCGGC
CGGCGGCCCTGACTGCGCTGCTGCTGCTGCTGCTGGGCCATGGCGGCGGCGGGCGCTGGGGCGCCCCGGGCCCAGG
AGGCGGCGGCGGCGGCGGCGGACGGGCCCCCGCGCGAGACGGCGAGGACGGACAGGACCCGACAGCAAGCACC
TGTACACGGCCGACATGTTTACGACCGGATCCAGAGCGCCGCGCACTTCGTTCATGTTCTTCGCGCCCTGGTGTG
GACACTGCCAGCGGCTGCAGCCGACTTGAATGACCTGGGAGACAATAACAGCAGTGAAGATGCCAAAGCTCT
ATGTGGCTAAAGTGGACTGCACGGCCCACTCCAGCTGTGCTCCGCCCAGGGGTGCGAGGATACCCCACTTAA
AGCTTTTCAAGCCAGGACCAAGAACTGTGAAGTACAGGGTCTCTCGGGACTTCCAGACACTGGAAAACCTGGATGC
TGCAGACACTGAACGAGGAGCCAGTGACACCAGAGCCGGAAAGTGGAACCGCCAGTGCCCCCGAGCTCAAGCAAG
GGCTGTATGAGCTCTCAGCAAGCAACTTTGAGCTGCACGTTGCACAAGGCGACCACTTTATCAAGTTCTTCGCTC
CGTGGTGTGGTCACTGCAAGCCCTGGCTCCAACCTGGGAGCAGCTGGCTCTGGGCTTGAACATTCCGAAACTG
TCAAGATTGGCAAGGTTGATTGTACACAGCACTATGAACCTGTCTCCGGAACACAGGTTCTGGGCTATCCCACCT
TTCTCTGTTTCCGAGATGGGAAAAAGTTGGATCAGTACAAGGGAAGCGGGATTTGGAGTCACTGAGGAGTACG
TGGAGTCGCGAGCTGCAGCGCACAGAGACTGGAGCGACGGAGACCGTCACGCCCTCAGAGGCCCCGGTGTGGCAG
CTGAGCCCGAGGCTGACAAGGGCACTGTGTTGGCACTCACTGAAAAATAACTTCGATGACCAATTGCAGAAGGAA
TAACCTTCATCAAGTTTTATGCTCCATGGTGTGGTCATTGTAAGACTCTGGCTCCTACTTTGGGAGGAACTCTCTA
AAAAGGAATTCCTGGTCTGGCGGGGGTCAAGATCGCCGAAGTAGACTGCAGTCTGAACGGAAATATCTGCAGCA
AGTATTCGGTACGAGGCTACCCACCGTTATTGCTTTTCCGAGGAGGGAAGAAAGTCAGTGAGCACAGTGAGGCA
GAGACCTTGACTCGTTACACCGCTTTGCTCTGAGCCAAGCGAAAGACGAACTTTAGGAACACAGTTGGAGGTCAC
CTCTCCTGCCAGCTCCCGCACCCCTGCGTTTAGGAGTTCACTCCACAGAGGCCACTGGGTTCCAGTGGTGGCT
GTTCAGAAAGCAGAACATACTAAGCGTGAGGTATCTTCTTTGTGTGTGTGTTTCCAAAGCCAACACACTCTACAG
ATTCTTTATTAAGTTAAGTTTCTCTAAGTAAATGTGTAACCTCATGGTCACTGTGTAACACTTTTCAGTGGCGATA
TATCCCCCTTTGACCTTCTCTTTGATGAAATTTACATGGTTTCCCTTTGAGACTAAACGTGTTAGGGGAAATGAAA
TTGCTGGACTATTTTGTGGCTCTGAGTTGAGTGATTTTGGTGAAAGGAAAGCATCCAAAGCATAGTTTACCTGC
CCACGAGTCTTGAAAGGGTGGCCTTGTGGCAGTATTGACGTTTCTCTGATCTTAAGGTCACAGTTGACTCAATAC
TGTGTTGGTCCGTAGCATGGAGCAGATTGAAATGCAAAAACCCACACCTCTGGAAGATACCTTCACGGCCGCTGC
TGGAGCTTCTGTTGCTGTGAATACTTCTCTCAGTGTGAGAGGTTAGCCGTGATGAAAGCAGCGTTACTTCTGACC
GTGCCTGAGTAAGAGAAATGCTGATGCCATAACTTTATGTGTGCGATACTTGTCAAATCAGTTACTGTTCAGGGGAT
CCTTCTGTTTCTCACGGGGTGAACACTGTCTTTAGTTTCTCATGTTTAACACGAAGCCAGGCCACATGAACTGT
TGGATGTTTCTTCTTAGAAAGGTAGGCATGGAATAATCCACGAGGCTCACTCTCAGTATCTCATTAACTCATTGA
AAGATTCCAGTTGTATTTGTACCTGGGGTGACAAGACCAGACAGGCTTTCCAGGCCCTGGGTATCCAGGGAGGC
TCTGCAGCCCTGCTGAAGGGCCCTAACTAGAGTTCTAGAGTTTCTGATTCTGTTTCTCAGTAGTCCTTTTAGAGG
CTTGCTATACCTGGTCTGCTTCAAGGAGGTCGACCTTCTAATGTTATGAAGATGGGATGCATTGTATCTCAAGAC
CAAAGACAGATGTCAGTGGGCTGCTCTGGCCCTGGTGTGACCGGCTGTGGCAGCTGTTGATGCCAGTGTCTCTTA
ACTCATGCTGTCTTGTGATTAAACACCTCTACTCTCCTTGGGAAATAAGCACATACAGGCTTAAAGCTCTAAGATA
GATAGGTGTTTTGCTCCTTTTACCATCGAGCTACTTCCCATAATAACCACTTTGTCATCCAACACTCTTCACCCACCT
CCCATAACGAAGGGGATGTGGATACTTGGCCCAAAGTAACTGGTGGTAGGAATCTTAGAAACAAGACCACCTTATA
CTGTCTGTCTGAGGCAGAAGATAACAGCAGCATCTCGACCAGCCTCTGCCTTAAAGGAAATCTTTATTAATCACG
TATGGTTCCAGATAAATCTTTTTTTTAAAAAAACCCAACTCCTAGAGAAGCACAACTGTCAAGAGTCTTGTACA
CAACAATTCAGCTTTGTCATCAGAGTCTGTATTCCAAGAAAATCAAAGTGGTACAATTTGTTTGTTTACACTAT
GATACTTTCTAAATAAACTCTTTTTTTTTTAA

FIGURE 35

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776

><subunit 1 of 1, 432 aa, 1 stop

><MW: 47629, pI: 5.90, NX(S/T): 0

MPARPGRLLPLLARPAALTALLLLLLGHGGGGRWGARAQEAAAAADGPPAADGEDGQDPHS
KHLYTADMFTHGIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAH
SDVCSAQGVRGYPTLKLFPKGQEAVKYQGPRDFQTLNWMQLTLNEEPVTPEPEVEPPSAPE
LKQGLYELSASNFEHVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT
QHYELCSGNQVRGYPTLLWFRDGGKVDQYKGKRDLESLREYVESQLQRTETGATETVTPSEA
PVLAAEPEADKGTVLALTENNFDDTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA
GVKIAEVDCTAERNICSKYSVRGYPTLLLFRRGGKKVSEHSGGRDLDSLHRFVLSQAKDEL

Signal sequence:

amino acids 1-32

[illegible]

CTTTTCTGAGGAACCAACAGCAATGAATGGCTTTGCATCCTTGCTTCGAAGAAACCAATTTAT
CCTCCTGGTACTATTTCTTTTGCAAATTCAGAGTCTGGGTCTGGATATTGATAGCCGTCCTA
CCGCTGAAGTCTGTGCCACACACAATTTCAACCAGGACCCAAAGGAGATGATGGTGAAAAA
GGAGATCCAGGAGAAGAGGGAAAGCATGGCAAAGTGGGACGCATGGGGCCGAAAGGAATTAA
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCCATTTGGGAAGA
AGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAAGCAGGTACT
GTCTGTGATTGTGGAAGATAACCGGAAATTTGTTGGACAACCTGGATATTAGTATTGCTCGGCT
CAAGACATCTATGAAGTTTGTCAAGAATGTGATAGCAGGGATTAGGGAAACTGAAGAGAAAT
TCTACTACATCGTGCAGGAAGAGAAGAACTACAGGGAATCCCTAACCCTGCAGGATTCGG
GGTGGAATGCTAGCCATGCCCAAGGATGAAGCTGCCAACACACTCATCGCTGACTATGTTGC
CAAGAGTGGCTTCTTTTCGGGTGTTTCATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA
TGTCACACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGGAACCCAGCGAC
CCCTATGGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG
CCATCTTACCATGTACTTTGTCTGTGAGTTCATCAAGAAGAAAAAGTAACTTCCCTCATCCT
ACGTATTTGCTATTTTTCCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTTTTTCCTG
ATTGTACTACATTTGATCTGAGTCAACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT
CCATCATCAAAAAAAAAAAAAAAAAA

FIGURE 37

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980

><subunit 1 of 1, 277 aa, 1 stop

><MW: 30645, pI: 7.47, NX(S/T): 2

MNGFASLLRRNQFILLVLFLLQIQSLGLDIDSRPTAEVCATHTISPGPKGDDGEKGDPGEEG
KHGKVGRMGPKGIKGELGDMGDQGNIGKTGPIGKKGDKGEKGLLGIPGEKKGKAGTVCDGCRY
RKFBVGQLDISIARLKTSMKFVKNVIAIGIRETEEFYIVQEEKNYRESLTHCRIRGGMLAMP
KDEAANTLIADYVAKSGFFRVFIGVNDLEREGQYMSTDNTPLQNYSNWNEGEPSDPYGHEDC
VEMLSSGRWNDTECHLTMYFVCEFIKKKK

Signal peptide:

amino acids 1-25

TOP SECRET

FIGURE 38

GGTTCTATCGATTCTGAATTCGGCCACACTGGCCGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCTGCTCTCCGCCCCGTGTGGAGTGGTGGGGCCCTGGGTGGGAATGGGCGTGT
GCCAGCGCACGCGCGCTCCCTGGAAGGAGAAAGTCTCAGCTAGAACGAGCGGCCCTAGGTTTT
CGGAAGGGAGGATCAGGGATGTTTGCGAGCGGCTGGAACCAGACGGTGCCGATAGAGGAAGC
GGGCTCCATGGCTGCCCTCCTGCTGCTGCCCCCTGCTGCTGTTGCTACCGCTGCTGCTGCTGA
AGCTACACCTCTGGCCGCGAGTTGCGCTGGCTTCCGGCGGACTTGGCCTTTGCGGTGCGAGCT
CTGTGCTGCAAAAGGGCTCTTCGAGCTCGCGCCCTGGCCGCGGCTGCCGCCGACCCGGAAGG
TCCCGAGGGGGGCTGCAGCCTGGCCTGGCGCCTCGCGGAAGTGGCCCAGCAGCGCGCCGCGC
ACACCTTTCTCATTACGGCTCGCGGCGCTTTAGCTACTCAGAGGCGGAGCGCGAGAGTAAC
AGGGCTGCACGCGCCTTCCCTACGTGCGCTAGGCTGGGACTGGGGACCCGACGGCGGCGACAG
CGGCGAGGGGAGCGCTGGAGAAGGCGAGCGGGCAGCGCCGGGAGCCGGAGATGCAGCGGCCG
GAAGCGGCGCGGAGTTTGCCGGAGGGGACGGTGCCGCCAGAGGTGGAGGAGCCGCCGCCCT
CTGTACCTGGAGCAACTGTGGCGCTGCTCCTCCCCGCTGGCCCAGAGTTTTCTGTGGCTCTG
GTTTCGGGCTGGCCAAGGCCGGCCTGCGCACTGCCTTTGTGCCACCGCCCTGCGCCGGGGCC
CCCTGCTGCACTGCCTCCGCAGCTGCGGCGCGCGCGCTGGTGCTGGCGCCAGAGTTTTCTG
GAGTCCCTGGAGCCGGACCTGCCCGCCCTGAGAGCCATGGGGCTCCACCTGTGGGCTGCAGG
CCCAGGAACCCACCTGCTGGAATTAGCGATTTGCTGGCTGAAGTGTCCGCTGAAGTGGATG
GGCCAGTGCCAGGATACCTCTCTCCCCCAGAGCATAACAGACACGTGCCTGTACATCTTC
ACCTCTGGCACCACGGGCCTCCCCAAGGCTGCTCGGATCAGTCATCTGAAGATCCTGCAATG
CCAGGGCTTCTATCAGCTGTGTGGTGTCCACCAGGAAGATGTGATCTACCTCGCCCTCCAC
TCTACCACATGTCCGGTTCCCTGCTGGGCATCGTGGGCTGCATGGGCATTGGGGCCACAGTG
GTGCTGAAATCCAAGTTCTCGGCTGGTCAAGTTCTGGGAAGATTGCCAGCAGCACAGGGTGAC
GGTGTTCAGTACATTGGGGAGCTGTGCCGATACCTTGTCAACCAGCCCCCGAGCAAGGCAG
AACGTGGCCATAAGGTCCGGCTGGCAGTGGGCAGCGGGCTGCGCCAGATACCTGGGAGCGT
TTTGTGCGGCGCTTCGGGCCCCCTGCAGGTGCTGGAGACATATGGACTGACAGAGGGCAACGT
GGCCACCATCAACTACACAGGACAGCGGGGCGCTGTGGGGCGTGCTTCTGGCTTTACAAGC
ATATCTTCCCCTTCTCCTTGATTGCTATGATGTCAACACAGGAGAGCCAATTCGGGACCCC
CAGGGGCACTGTATGGCCACATCTCCAGGTGAGCCAGGGCTGCTGGTGGCCCCGGTAAGCCA
GCAGTCCCCATTCTGGGCTATGCTGGCGGGCCAGAGCTGGCCCAGGGGAAGTTGCTAAAGG
ATGTCTTCCGGCCTGGGGATGTTTTCTTCAACACTGGGGACCTGCTGGTCTGCGATGACCAA
GGTTTTCTCCGCTTCCATGATCGTACTGGAGACACCTTCAGGTGGAAGGGGGAGAATGTGGC
CACAACCGAGGTGGCAGAGGTCTTCGAGGCCCTAGATTTTCTTCAGGAGGTGAACGTCTATG
GAGTCACTGTGCCAGGGCATGAAGGCAGGGCTGGAATGGCAGCCCTAGTTCTGCGTCCCCC
CACGCTTTGGACCTTATGCAGCTCTACACCCACGTGTCTGAGAACTTGCCACCTTATGCCCG
GCCCCGATTCTCAGGCTCCAGGAGTCTTTGGCCACCACAGAGACCTTCAAACAGCAGAAAG
TTCGGATGGCAAATGAGGGCTTCGACCCACAGCACCTGTCTGACCCACTGTACGTTCTGGAC
CAGGCTGTAGGTGCCTACCTGCCCTCACAAGTCCCGGTACAGCGCCCTCCTGGCAGGAAA
CCTTCGAATCTGAGAACTTCCACACCTGAGGCACCTGAGAGAGGAACTCTGTGGGGTGGGG
CCGTTGCAGGTGTACTGGGCTGTGAGGATCTTTTCTATAACAGAACTGCGGTCACTATTTT
GTAATAAATGTGGCTGGAGCTGATCCAGCTGTCTCTGACCTAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAGGCGGCCGCGACTCTAGAGTCGACCTGCAGTAGGGATAACAGGGTAATAAGC
TTGGCCGCCATGGCCCACTTGTATTATTGAG

FIGURE 39

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913

><subunit 1 of 1, 730 aa, 1 stop

><MW: 78644, pI: 7.65, NX(S/T): 2

MGVCQRTAPWKEKSQLERAAALGFRKGGSGMFASGWNQTVPIEEAGSMAALLLLPLLLLLPL
LLKLHLWPQLRWLPADLAFVRLCCKRALRARALAAAAADPEGPEGGCSLAWRLAELAQQ
RAAHTFLIHGSRRFSYSEAERESNRAARAFLRALGWDWGPDPGDSGEGSAGEGERAAPGAGD
AAAGSGAEFAGGDGAARGGGAAAPLSPGATVALLLPAGPEFLWLWFGGLAKAGLRTAFVPTAL
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA
EVDGPVPGYLSSPQSITDTCLYIFTSGTTGLPKAARISHLKILQCQGFYQLCGVHQEDVIYL
ALPLYHMSGSLLGIVGCMGIGATVVLKSKFSAGQFWEDCQQHRVTVFQYIGELCRYLVNQPP
SKAERGHKVPRLAVGSGLRPDTWERFVRRFGPLQVLETYGLTEGNVATINYTGORGAVGRASW
LYKHIFPFSLIRYDVTTGEPIRDPOGHCMATSPGEPGLLVAPVSQQSPFLGYAGGPPELAQKG
LLKDVFVRPGDVFFNTGDLLVCDDQGFRLFHDRTGDTFRWKGENVATTEVAEVFEALDFLQEV
NVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLPYARPRFLRLQESLATTETFK
QQKVRMANEGFDPSTLSDPLYVLDQAVGAYLPLTTARYSALLAGNLRI

Type II transmembrane domain:

amino acids 45-65

Other transmembrane domain:

amino acids 379-398

cAMP- and cGMP-dependent protein kinase phosphorylation site
starting at amino acid 136

CUB domain protein motif

amino acids 254-261

putative AMP-binding domain signature

amino acids 332-343

N-glycosylation sites

amino acids 37-40 and 483-486

FIGURE 40

CCTGTGTTAAGCTGAGGTTTCCCCTAGATCTCGTATATCCCCAACACATACCTCCACGCACA
CACATCCCCAAGAACCTCGAGCTCACACCAACAGACACACGCGCGCATACACACTCGCTCTC
GCTTGTCCATCTCCCTCCCGGGGAGCCGGCGCGCTCCCACCTTTGCCGCACACTCCGGC
GAGCCGAGCCCGCAGCGCTCCAGGATTCTGCGGCTCGGAACTCGGATTGCAGCTCTGAACCC
CCATGGTGGTTTTTTTAAACACTTCTTTTCTTCTCTTCTCGTTTTTGATTGCACCGTTTTCCA
TCTGGGGGCTAGAGGAGCAAGGCAGCAGCCTTCCCAGCCAGCCCTTGTTGGCTTGCCATCGT
CCATCTGGCTTATAAAAGTTTTGCTGAGCGCAGTCCAGAGGGCTGCGCTGCTCGTCCCCCTCGG
CTGGCAGAAGGGGGTGACGCTGGGCAGCGCGAGGAGCGCGCCGCTGCCTCTGGCGGGCTTT
CGGCTTGAGGGGCAAGGTGAAGAGCGCACCGGCCGTGGGGTTTTACCGAGCTGGATTTGTATG
TTGCACCATGCCCTTCTTGATCGGGGCTGTGATTCTTCCCCCTCTTGGGGCTGCTGCTCTCCC
TCCCCGCCGGGGCGGATGTGAAGGCTCGGAGCTGCGGAGAGGTCCGCCAGGCGTACGGTGCC
AAGGGATTACGCCTGCGGACATCCCCTACCAGGAGATCGCAGGGGAACACTTAAGAATCTG
TCCTCAGGAATATACATGCTGCACCACAGAAATGGAAGACAAGTTAAGCCAACAAAGCAAAC
TCGAATTTGAAAACCTTGTGGAAGAGACAAGCCATTTTGTGCGCACCACTTTTGTGTCCAGG
CATAAGAAATTTGACGAATTTTTCCGAGAGCTCCTGGAGAATGCAGAAAAGTCACTAAATGA
TATGTTTTGTACGGACCTATGGCATGCTGTACATGCAGAATTCAGAAGTCTTCCAGGACCTCT
TCACAGAGCTGAAAAGGTACTACACTGGGGGTAATGTGAATCTGGAGGAAATGCTCAATGAC
TTTTGGGCTCGGCTCCTGGAACGGATGTTTCAGCTGATAAACCTCAGTATCACTTCAGTGA
AGACTACCTGGAATGTGTGAGCAAATACACTGACCAGCTCAAGCCATTTGGAGACGTGCCCC
GGAACTGAAGATTCAGGTTACCCGCGCCTTCATTGCTGCCAGGACCTTTGTCCAGGGGCTG
ACTGTGGGCAGAGAAGTTGCAAACCGAGTTTCCAAGGTCAGCCCAACCCAGGGTGTATCCG
TGCCCTCATGAAGATGCTGTACTGCCATACTGTGCGGGGCTTCCCCTGTGAGGCCCTGCA
ACAATACTGTCTCAACGTATGAAGGGCTGCTTGGCAAATCAGGCTGACCTCGACACAGAG
TGGAATCTGTTTTATAGATGCAATGCTCTTGGTGGCAGAGCGACTGGAGGGGCCATTCAACAT
TGAGTCGGTCATGGACCCGATAGATGTCAAGATTTCTGAAGCCATTATGAACATGCAAGAAA
ACAGCATGCAGGTGTCTGCAAAGGTCTTTTCAGGGATGTGGTCAGCCCAACCTGCTCCAGCC
CTCAGATCTGCCCCGCTCAGCTCCTGAAAATTTTAATACAGTTTTCAGGCCCTACAATCCTGA
GGAAAGACCAACAACCTGCTGCAGGCACAAGCTTGGAACCGGCTGGTCACAGACATAAAAGAGA
AATTGAAGCTCTCTAAAAGGTCTGGTCAGCATTACCCTACACTATCTGCAAGGACGAGAGC
GTGACAGCGGGCACGTCCAACGAGGAGGAATGCTGGAACGGGCACAGCAAAGCCAGATACTT
GCCTGAGATCATGAATGATGGGCTCACCAACCAGATCAACAATCCCGAGGTGGATGTGGACA
TCACTCGGCCTGACACTTTTCATCAGACAGCAGATTATGGCTCTCCGTGTGATGACCAACAAA
CTAAAAAACGCCTACAATGGCAATGATGTCAATTTCCAGGACACAAGTGATGAATCCAGTGG
CTCAGGGAGTGGCAGTGGGTGCATGGATGACGTGTGTCCACGGAGTTTGAGTTTGTACCA
CAGAGGCCCCCGCAGTGGATCCCGACCGGAGAGAGGTGGACTCTTCTGCAGCCCAGCGTGCC
CACTCCCTGCTCTCCTGGTCTCTCACCTGCATTGTCCTGGCACTGCAGAGACTGTGCAGATA
ATCTTGGGTTTTTTGGTCAGATGAACTGCATTTTAGCTATCTGAATGGCCAACTCACTTCTT
TTCTTACACTCTTGGACAATGGACCATGCCACAAAACTTACCGTTTTCTATGAGAAGAGAG
CAGTAATGCAATCTGCCTCCCTTTTTGTTTTCCCAAAGAGTACCGGTGCCAGACTGAACTG
CTTCTCTTTCCTTCAGCTATCTGTGGGGACCTTGTTTATTCTAGAGAGAATTCTTACTCAA
ATTTTTTCGTACCAGGAGATTTTCTTACCTTCATTTGCTTTTATGCTGCAGAAGTAAAGGAAT
CTCACGTTGTGAGGGTTTTTTTTTTCTCATTTAAAT

FIGURE 41

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914

><subunit 1 of 1, 555 aa, 1 stop

><MW: 62736, pI: 5.36, NX(S/T): 0

MPSWIGAVILPLLGLLLSLPAGADV KARS CGEVRQAYGAKGFSLADIPYQEIAGEHLRICPQ
EYTCCTTEMEDKLSQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLEN A E K S L N D M F
VRTYGM L Y M Q N S E V F Q D L F T E L K R Y Y T G G N V N L E E M L N D F W A R L L E R M F Q L I N P Q Y H F S E D Y
L E C V S K Y T D Q L K P F G D V P R K L K I Q V T R A F I A A R T F V Q G L T V G R E V A N R V S K V S P T P G C I R A L
M K M L Y C P Y C R G L P T V R P C N N Y C L N V M K G C L A N Q A D L D T E W N L F I D A M L L V A E R L E G P F N I E S
V M D P I D V K I S E A I M N M Q E N S M Q V S A K V F Q G C G Q P K P A P A L R S A R S A P E N F N T R F R P Y N P E E R
P T T A A G T S L D R L V T D I K E K L K L S K K V S A L P Y T I C K D E S V T A G T S N E E E C W N G H S K A R Y L P E
I M N D G L T N Q I N N P E V D V D I T R P D T F I R Q Q I M A L R V M T N K L K N A Y N G N D V N F Q D T S D E S S G S G
S G S G C M D D V C P T E F E F V T T E A P A V D P D R R E V D S S A A Q R G H S L L S W S L T C I V L A L Q R L C R

Signal peptide:

amino acids 1-23

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 42A

CGGACGCGTGGGCGGACGCGTGGGCAAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAACGCACG
CAGTTTGCAGCGCCTGCGCCGGGTGCGCCAACTACGCAAAGACCAAGCGGGCTCCGCGCGGACCGGCCGCGGGG
TAGGGACCCGGCTTTGGCCTTCAGGCTCCCTAGCAGCGGGGAAAAGGAATTGCTGCCCGGAGTTTCTGCGGAGGT
GGAGGGAGATCAGGAAACGGCTTCTTCTCACTTCGCGCCCTGGTGTAGTGTGCGGGGAGATTGGCAAACGCCTAGG
AAAGGACTGGGGAAAATAGCCCTGGGAAAAGTGGAGAAGGTGATCAGGAGGCCGGTCCACTACGGCAGTTTATCTG
TCTGATCAGAGCCAGACGCGACGCGTCCACTTCGCAGTTCTTTCCAGGTGTGGGGACCGCAGGACAGACGGCCGA
TCCCCGCCGCCCTCCGTACCAGCACTCCCAGGAGAGTCAGCCTCGCTCCCCAACGTGAGGGCGCTCTGGCCACGA
AAAGTTCCTGTCCACTGTGATTCTCAATTCTTGCTTGGTTTCTTCTCCAGAGAACTTTTGGGTGGAGATATTA
ACTTTTTTCTTTTTTTTTTCTTGGTGGAAAGCTGCTCTAGGGAGGGGGAGGAGGAGGAGAAAGTGAAATGTGC
TGGAGAAGAGCGAGCCCTCCTTGTCTTCCGGAGTCCCATCCATTAAGCCATCACTTCTGGAAGATTAAAGTTGT
CGGACATGGTGACAGCTGAGAGGAGAGGAGGATTTCTTGCCAGGTGGAGAGTCTTACCCTGTGTGGGTGCATG
TGTGCGCCCGCAGCGGCGCGGGGCGCGTGGTTCTCCGCTGGAGTCTCACCTGGGACCTGAGTGAATGGCTCCCA
GGGGCTGTGCGGGGCATCCGCCTCCGCCTTCTCCACAGGCCTGTGTCTGTCTGGAAAGATGCTAGCAATGGGGG
CGCTGGCAGGATTCTGGATCCTCTGCCCTCCTCACTTATGGTTACCTGTCTTGGGGCCAGGCCTTAGAAGAGGAGG
AAGAAGGGGCCTTACTAGCTCAAGCTGGAGAGAAAAGTACAGCCAGCACAACCTCCACCTCCCAGCCCCATCTCA
TTTTCATCCTAGCGGATGATCAGGGATTAGAGATGTGGGTACCACGGATCTGAGATTAAACACCTACTCTTG
ACAAGCTCGCTGCCGAAGGAGTTAAACTGGAGAAGTACTATGTCCAGCCTATTTGCACACCATCCAGGAGTCAGT
TTATTACTGGAAGTATCAGATACACACCGGACTTCAACATTTCTATCATAAGACCTACCCAACCCAACTGTTTAC
CTCTGGACAATGCCACCCTACCTCAGAAAAGTGAAGGAGGTGGATATTCAACGCATATGGTCGGAAAATGGCACT
TGGGTTTTTAACAGAAAAGAAATGCATGCCACCAGAAAGAGGATTTGATACCTTTTTTGGTTCCCTTTTGGGAAGTG
GGGATTACTATACACACTACAAATGTGACAGTCCCTGGGATGTGTGGCTATGACTTGTATGAAAACGACAATGCTG
CCTGGGACTATGACAATGGCATATACTCCACACAGATGTACACTCAGAGAGTACAGCAAATCTTAGCTTCCATA
ACCCACAAAGCCTATATTTTTATATACTGCCTATCAAGCTGTTTATTACCAGTGAAGCTCCTGGCAGGTATT
TCGAACACTACCGATCCATTATCAACATAAACAGGAGAAGATATGCTGCCATGCTTTCCTGCTTAGATGAAGCAA
TCAACAACGTGACATTGGCTCTAAAGACTTATGGTTTCTATAACAACAGCATTATCATTTACTCTTCAGATAATG
GTGGCCAGCCTACGGCAGGAGGGAGTAACTGGCCTCTCAGAGGTAGCAAAGGAACATATTGGGAAGGAGGGATCC
GGGCTGTAGGCTTTGTGCATAGCCCACTTCTGAAAAACAAGGGAACAGTGTGTAAGGAACCTGTGCACATCACTG
ACTGGTACCCCACTCTCATTTCACTGGCTGAAGGACAGATTGATGAGGACATTCAACTAGATGGCTATGATATCT
GGGAGACCATAAGTGAGGGTCTTCGCTCACCCCGAGTAGATATTTGCATAACATTGACCCCTATACACCAAGGC
AAAAAATGGCTCCTGGGCAGCAGGCTATGGGATCTGGAACACTGCAATCCAGTCAGCCATCAGAGTGCAGCACTG
GAAATTGCTTACAGGAAATCCTGGCTACAGCGACTGGGTCCCCCTCAGTCTTTCAGCAACCTGGGACCGAACC
GTGGCACAATGAACGGATCACCTTGTCAACTGGCAAAAGTGTATGGCTTTTCAACATCACAGCCGACCCATATGA
GAGGGTGGACCTATCTAACAGGTATCCAGGAATCGTGAAGAAGCTCCTACGGAGGCTCTCACAGTTCAACAAAAC
TGCAGTGCCGCTCAGGTATCCCCCAAAGACCCAGAAGTAACCTAGGCTCAATGGAGGGGTCTGGGGACCATG
GTATAAAGAGGAAACCAAGAAAAAGAAAGCAAGCAAAAATCAGGCTGAGAAAAAGCAAAAGAAAAGCAAAAAA
GAAGAAGAAACAGCAGAAAGCAGTCTCAGGTAAACCAGCAAATTTGGCTCGATAATATCGCTGGCCTAAGCGTCA
GGCTTGTTTTCATGCTGTGCCACTCCAGAGACTTCTGCCACCTGGCCGCCACACTGAAAAGTGTCTGCTCAGTG
CCAAGGTGCTACTCTTGCAAGCCACACTTAGAGAGAGTGGAGATGTTTATTTCTCTCGCTCCTTTAGAAAACGTG
GTGAGTCTGAGTTCCTGCTGTGCTTCAGTCAACTGACCAAAACACTGCTTTGAATTATAGGAGGAGAAACAATA
ACCTACCATCCGCAAGCATGCTAATTTGATGGAAGTTACAGGGTAGCATGATTAAACCTACCTTTGATAAATTAC

FIGURE 42B

AGTCAAAGATTGTGTCACCTCAAAGGCCTTGAAGAATATATTTTCTTGGTGAATTTTGTATGTCTGTCATATGA
CACTTGGGTTTTTTAATTAAATCTATTTTATATATATAAATATATGTTTCTTTTCTGTGAAAAGCTGTTTTCT
CACATGTGAACAGCTTGCACCTCATTTTACCATGCGTGAGGGAATGGCAAATAAGAATGTTTGAGCACACTGCCC
ACAATGAATGTAACATTTTTCTAAACACTTTACTAGAAGAACATTTTCAGTATAAAAAACCTAATTTATTTTTACA
GAAAAATATTTTGTGTGTTTTTATAAAAAAGTTATGCAAATGACTTTTATTTTTATTTCTGCATACCATTAGAAGA
ATTTTATTTTCATTTCTTCAAATTATCAAGCACTGTAATACTATAAATTAATGTAATACTGTGTGAATTCAGACTA
TAAAAAACATCATTCAGAAAACTTTATAATCGTCATTGTTCAATCAAGATTTTGAATGTAATAAGATGAATATAT
ATTACTTGGAATTCATGTTTGTGTCAGAGTTGAGACAACCTTTATTGTTTCTATCATAAACTATTTATGTATCTT
AATTATTAAAAATGATTTACTTTATGGCACTAGAAAATTTACTGTGGCTTTTCTGATCTAACTTCTAGCTAAAAT
GTATCATTTGGTCCATAAAAAATAAAATCTTTACTAATAGGCAATTGAAGGAATGGTTTGCTAACCAACCACAGTAA
TATAATATGATTTTACAGATAGATGCTTCCCCTTGGCTATGACATGGAGAAAGATTTTCCATAATAATAACTAA
TATTTATATTAGGTTGGTGCAAACTAGTTGCGGTTTTTCCCATTAAGTAATAACCTTACTCTTATACAAAGT
GGACACTGTGGGGAGATACAGAGAAATGGAAGATACGGATCCTGCCTGGAGTAGGTAACCTTGCTTGGAACCCC
ACATGCAAACGTCATGAGGAGAATTAAAGGAGTATTATCAGTAATGAAGTTTATCATGGGTCATCAATGAGCATA
GATTGGTGTGGATCCTGTAGACCCTGGTGTTTTCTTTGAAGTGCCCTCTCCTAATGCAGAGGCCTTGAAGCTTAC
AGTATACACTTGAAAAGTCACAGATAGCTAGAATTATGATCTTTGAAGTTATAACTGTGATCTGAAAATGTGTGT
GGTGGTATGACAGCATACCATTAAATACATTTACATCACAGCTCAAAGGACTGTGATATAATCCATTTATATCAC
AACTCAAAGGACTGTGATATAATCCATTTATATCACAGCTCACAGTTTCTGAAAATGTATAAAAGAATCTATAAT
CTAGTACTGAAATTACTAAATTGGGTAAGATGATTTAAATGATTTTAATTTTAACATTTTATTTCTAGAATATAT
GGCTCCATTTTATTTTATAGTGTAAGTTGTATTTCTTAAAGTTTGTGTTTTGTGCGACAGTATCTTTTAAATGAG
TCTTAAAAATAAAGGCATATTGTTTCATGTTTAAA
AAA

FIGURE 43

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296

><subunit 1 of 1, 515 aa, 1 stop

><MW: 56885, pI: 6.49, NX(S/T): 5

MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTGYLSWGQALEEEEEEGALLAQA
GEKLEPSTTSTSQPHLIFILADDQGFRDVGYPHGSEIKTPTLDKLAAGVKLENYYVQPICTP
SRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMGKWHLGFNKKEC
MPTRRGFDTFFGSLLGSGDYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQ
ILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLEAINNVTLA
LKTYGFYNNSIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGVHSPLLKNKGTVCK
ELVHITDWYPTLISLAEGQIDEDIQLDGDIWETISEGLRSPRVDILHNIDPYTPRQKMAPG
QQAMGSGTLQSSQPSECSTGNCLQEILATATGSPLSLSATWDRTGGTMNGSPCQLAKVYGFS
TSQPTHMRGWTYLTGIQES

Important Features:

Signal Peptide:

amino acids 1-37

Sulfatases signature 1.

amino acids 120-132

Sulfatases signature 2.

amino acids 168-177

Tyrosine kinase phosphorylation site.

amino acids 163-169

N-glycosylation sites.

amino acids 157-160, 306-309 and 318-321

FIGURE 44

CGGACGCGTGGGTGCGAGTGGAGCGGAGGACCCGAGCGGCTGAGGAGAGAGGAGGCGGCGGC
TTAGCTGCTACGGGGTCCGGCCGGCGCCCTCCCGAGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGAGAATGCCTCTGCCCTGGAGCCTTGCGCTCCCGCTGCTGCTCTCCTGGGTGGCAG
GTGGTTTCGGGAACGCGGCCAGTGCAAGGCATCACGGGTTGTTAGCATCGGCACGTCAGCCT
GGGTCTGTCACTATGGAACATAAAGTGGCCTGCTGCTACGGCTGGAGAAGAAACAGCAAGGG
AGTCTGTGAAGCTACATGCGAACCTGGATGTAAGTTTGGTGAGTGCGTGGGACCAACAAAT
GCAGATGCTTTCCAGGATACACCGGGAAACCTGCAGTCAAGATGTGAATGAGTGTGGAATG
AAACCCCGGCCATGCCAACACAGATGTGTGAATACACACGGAAGCTACAAGTGCTTTTGCCT
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGTGAACCTCTAGGACATGTGCCATGATAA
ACTGTCACTACAGCTGTGAAGACACAGAAGAAGGGCCACAGTGCCTGTGTCCATCCTCAGGA
CTCCGCCTGGCCCCAAATGGAAGAGACTGTCTAGATATTGATGAATGTGCCTCTGGTAAAGT
CATCTGTCCCTACAATCGAAGATGTGTGAACACATTTGGAAGCTACTACTGCAAAATGTCACA
TTGGTTTCGAAGTGAATATATCAGTGGACGATATGACTGTATAGATATAAATGAATGTACT
ATGGATAGCCATACGTGCAGCCACCATGCCAATTGCTTCAATACCCAAGGGTCCTTCAAGTG
TAAATGCAAGCAGGGATATAAAGGCAATGGACTTCGGTGTTCTGCTATCCCTGAAAATTCTG
TGAAGGAAGTCTCAGAGCACCTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCTCAC
AAAAACAGCATGAAAAAGAAGGCAAAAATTAAAAATGTTACCCAGAACCCACCAGGACTCC
TACCCCTAAGGTGAACCTGCAGCCCTTCAACTATGAAGAGATAGTTTCCAGAGGCGGGAAC
CTCATGGAGGTAAAAAAGGGAATGAAGAGAAATGAAGAGGGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCCTGAAGAATGACATAGAGGAGCGAAGCCTGCGAGGAGATGTGTTTTTCCCT
AAGGTGAATGAAGCAGGTGAATTTCGGCCTGATTCTGGTCCAAAGGAAAGCGCTAACTTCCAA
ACTGGAACATAAAGATTTAAATATCTCGGTTGACTGCAGCTTCAATCATGGGATCTGTGACT
GGAAACAGGATAGAGAAGATGATTTTGAAGTGAATCCTGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTTCCGGCCTTGGCAGGTACAAAGAAAGACATTGGCCGATTGAAACTTCT
CCTACCTGACCTGCAACCCCAAAGCAACTTCTGTTTGCTCTTTGATTACCGGCTGGCCGGAG
ACAAAGTCGGGAAACTTCGAGTGTTTGTGAAAAACAGTAACAATGCCCTGGCATGGGAGAAG
ACCACGAGTGAGGATGAAAAGTGAAGACAGGGAAAATTGAGTTGTATCAAGGAACTGATGC
TACCAAAAGCATCATTTTTGAAGCAGAACGTGGCAAGGGCAAAACCGGCGAAATCGCAGTGG
ATGGCGTCTTGCTTGTTTCAGGCTTATGTCCAGATAGCCTTTTATCTGTGGATGACTGAATG
TTACTATCTTTATATTGACTTTGTATGTCAGTTCCTTGTTTTTTTGTATATTGCATCATAG
GACCTCTGGCATTTTGAATTACTAGCTGAAAAATTGTAATGTACCAACAGAAATATTATTG
TAAGATGCCTTTCTTGATATAAGATATGCCAATATTTGCTTTAAATATCATATCACTGTATCT
TCTCAGTCATTTCTGAATCTTTCNCATTATATTATAAAATNTGGAAANGTCAGTTTATCTC
CCCTCCTCNGTATATCTGATTTGTATANGTANGTTGATGNGCTTCTCTCTACAACATTTCTA
GAAAATAGAAAAAAAAGCACAGAGAAATGTTTAACTGTTTGACTCTTATGATACTTCTTGGA
AACTATGACATCAAAGATAGACTTTTGCCTAAGTGGCTTAGCTGGGTCTTTCATAGCCAAAC
TTGTATATTTAATTCTTTGTAATAATAA

FIGURE 45

MPLPWSLALPLLLSWVAGGFGNAASARHHGLLASARQPGVCHYGTKLACCYGWRNSKGVCE
ATCEPGCKFGECVGPKNKRCFPGYTGKTCSQDVNECGMKPRPCQHRCVNTHGSKCFCLSGH
MLMPDATCVNSRTCAMINCQYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDECASGKVICP
YNRRCVNTFGSYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFNTQGSFKCKCK
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKLLAHKNSMKKKAKIKNVTPEPTRTPTPK
VNLQPFNYEEIVSRGGNSHGGKKGNEEK

Signal peptide:

amino acids 1-21

EGF-like domain cysteine pattern signature.

amino acids 80-91

Calcium-binding EGF-like domains

amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGTGGCTGCTGGTGCTGTGCGCGCTGCTCCTGCTCTTGGTGCAGCTGCTG
CGCTTCCTGAGGGCTGACGGCGACCTGACGCTACTATGGGCCGAGTGGCAGGGACGACGCCC
AGAATGGGAGCTGACTGATATGGTGGTGTGGGTGACTGGAGCCTCGAGTGGAATTGGTGAGG
AGCTGGCTTACCAGTTGTCTAAACTAGGAGTTTCTCTTGCTGCTGTCAGCCAGAAGAGTGCAT
GAGCTGGAAAGGGTGAAAAGAAGATGCCTAGAGAATGGCAATTTAAAAGAAAAAGATATACT
TGTTTTGCCCTTGACCTGACCGACACTGGTTCCCATGAAGCGGCTACCAAAGCTGTTCTCC
AGGAGTTTGGTAGAATCGACATTCTGGTCAACAATGGTGGAATGTCCAGCGTTCTCTGTGC
ATGGATAACCAGCTTGGAATGTCTACAGAAAGCTAATAGAGCTTAACTACTTAGGGACGGTGTC
CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA
ATAGCATCCTGGGTATCATATCTGTACCTCTTTCCATTGGATACTGTGCTAGCAAGCATGCT
CTCCGGGGTTTTTTTTAATGGCCTTCGAACAGAACTTGCCACATACCCAGGTATAATAGTTTC
TAACATTTGCCCAGGACCTGTGCAATCAAATATTGTGGAGAATTCCCTAGCTGGAGAAGTCA
CAAAGACTATAGGCAATAATGGAGACCAGTCCCACAAGATGACAACCAGTCGTTGTGTGCGG
CTGATGTTAATCAGCATGGCCAATGATTTGAAAGAAGTTTGGATCTCAGAACAACCTTTCTT
GTTAGTAACATATTTGTGGCAATACATGCCAACCTGGGCCTGGTGGATAACCAACAAGATGG
GGAAGAAAAGGATTGAGAACTTTAAGAGTGGTGTGGATGCAGACTCTTCTTATTTTAAAATC
TTTAAGACAAAACATTGACTGAAAAGAGCACCTGTACTTTTCAAGCCACTGGAGGGAGAAATG
GAAAACATGAAAACAGCAATCTTCTTATGCTTCTGAATAATCAAAGACTAATTTGTGATTTT
ACTTTTTAATAGATATGACTTTGCTTCCAACATGGAATGAAATAAAAAATAAATAATAAAG
ATTGCCATGAATCTTGCAAAA

FIGURE 47

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343

><subunit 1 of 1, 289 aa, 1 stop

><MW: 32268, pI: 9.21, NX(S/T): 0

MVVWVTGASSGIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLENGNLKEKDILVLPDL
TDTGSHEAATKAVLQEFGRIDILVNNGGMSQRSLCMDTSLDVYRKLIELNYLGTVSLTKCVL
PHMIERKQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIIVSNICPGP
VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSCVRLMLISMANDLKEVWISEQPFLLVTYLW
QYMPWTAWWITNKMGGKRIENFKSGVDADSSYFKIFKTKHD

Important Features:

Signal Peptide:

amino acids 1-31

Transmembrane domain:

amino acids 136-157

Tyrosine kinase phosphorylation site.

106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase

amino acids 80-90, 131-168, 1-13 and 176-185

FIGURE 48

GCGACGTGGGCACCGCCATCAGCTGTTTCGCGCGTCTTCTCCTCCAGGTGGGGCAGGGGTTTC
GGGCTGGTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCCAACAGCATATTCGG
TTGCATCTTCTACACACTACAGCTATTGTTAGGTTGCCTGCGGACACGCTGGGCCTCTGTCC
TGATGCTGCTGAGCTCCCTGGTGTCTCTCGCTGGTTCTGTCTACCTGGCCTGGATCCTGTTT
TTCGTGCTCTATGATTTCTGCATTGTTTGTATCACCACCTATGCTATCAACGTGAGCCTGAT
GTGGCTCAGTTTCCGGAAGGTCCAAGAACCCAGGGCAAGGCTAAGAGGCACTGAGCCCTCA
ACCCAAGCCAGGCTGACCTCATCTGCTTTTGTCTTTGGTCTTCAAGCCGCTCAGCGTGCCTGTG
GACAGCGTGGCCCCCGCCCCCAAGCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCCC
TGGCAGGCCAGTGTGAGGAGGCAAGGAGCCCACATCTGCAGCGGCTCCCTGGTGGCAGACAC
CTGGGTCTCACTGCTGCCCACTGCTTTGAAAAGGCAGCAGCAACAGAACTGAATTCTGGT
CAGTGGTCTGGGTTCTCTGCAGCGTGAGGGACTCAGCCCTGGGGCCGAAGAGTGGGGGTG
GCTGCCCTGCAGTTGCCAGGGCCTATAACCACTACAGCCAGGGCTCAGACCTGGCCCTGCT
GCAGCTCGCCACCCACGACCCACACACCCCTCTGCCTGCCCCAGCCCGCCCATCGCTTCC
CCTTTGGAGCCTCCTGCTGGGCCACTGGCTGGGATCAGGACACCAGTGATGCTCCTGGGACC
CTACGCAATCTGCGCCTGCGTCTCATCAGTCGCCCCACATGTAAGTGTATCTACAACCAGCT
GCACCAGCGACACCTGTCCAACCCGGCCCCGGCCTGGGATGCTATGTGGGGGCCCCCAGCCTG
GGGTGCAGGGCCCCCTGTGAGGGAGATTCCGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGA
CACTGGGTTTCAAGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCCAGGAGGACGCTCCTGT
GCTGCTGACCAACACAGCTGCTCACAGTTCTGGCTGCAGGCTCGAGTTCAGGGGGCAGCTT
TCCTGGCCCAGAGCCCAGAGACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGA
TCCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCTCCCCATGGCCCTGGGAGGCCAGGCT
GATGCACCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTGAGAGGAGGCGGTGCTAACTG
CTGCCCCTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCAGA
CCGGAGGAGTGGGGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCCACCCTGAGGGGGG
CTACGACATGGCCCTCCTGCTGCTGGCCCAGCCTGTGACACTGGGAGCCAGCCTGCGGCCCC
TCTGCCTGCCCTATCCTGACCACCACCTGCCTGATGGGGAGCGTGGCTGGGTTCTGGGACGG
GCCCCGCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCCTCCTGGGGCCTAG
GGCCTGCAGCCGGCTGCATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGG
TGTGTACCAAGTGCTGTGGGTGAGCTGCCAGCTGTGAGGGCCTGTCTGGGGCACCCTGGTG
CATGAGGTGAGGGGCACATGGTTCTGGCCGGGCTGCACAGCTTCGGAGATGCTTGCCAAGG
CCCCGCCAGGCCGGCGGTCTTACCAGCGCTCCCTGCCTATGAGGACTGGGTGAGCAGTTTGG
ACTGGCAGGTCTACTTTCGCCGAGGAACCAAGAGCCCGAGGCTGAGCCTGGAAGCTGCCTGGCC
AACATAAGCCAACCAACCAGCTGCTTGACAGGGGACCTGGCCATTCTCAGGACAAGAGAATGC
AGGCAGGCAAATGGCATTACTGCCCCCTGTCTCCCCACCCTGTATGTGTGATTCCAGGCAC
CAGGGCAGGCCCAGAAGCCCAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGGTGCCCCA
CTCCCCACCCTGCAGGACAGGGGTGTCTGTGGACACTCCACACCCAACTCTGCTACCAAGC
AGGCGTCTCAGCTTTCCTCCTCCTTTACTCTTTCAGATACAATCACGCCAGCCACGTTGTTT
TGAAAATTTCTTTTTTTGGGGGGCAGCAGTTTTCTTTTTTTAAACTTAAATAAATTGTTAC
AAAATAAAA

FIGURE 49

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571

MLLSSLVSLAGSVYLAWILFFVLYDFCIVCITTYAINVSLMWLSFRKVQEPQGKAKRHGNTV
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSDDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTS
DAPGTLRNLRLRLISRPTCNCIYNQLHQRHLSNPARPGMLCGGPQPGVQGPCQGDSGGPVLC
LEPDGHWVQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAAFLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQQLACGGALVSEEAVLTAAHCFIGRQAPEEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYPDHHLDPDGERG
WVLGRARPGAGISSLQTPVTLLGPRACSRLLHAAPGGDGSPILPGMVCTSAVGELPSCEGLS
GAPLVHEVRGTWFLAGLHSFGDACQGPAPAVFTALPAYEDWVSSLDWQVYFAEPEPEPEAEP
GSCLANISQPTSC

Important features:

Signal peptide:

amino acids 1-15

Homologous region to Serine proteases, trypsin family

amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.

amino acids 37-40 and 564-567

Kringle domains

amino acids 79-96, 343-360 and 235-247

FIGURE 50

CGGGCCGCCCCCGGCCCCATTTCGGGCGGGCCTCGCTGCGGCGGCGACTGAGCCAGGCTGG
GCCGCGTCCCTGAGTCCCAGAGTCGGCGCGGCGCGGCAGGGGCAGCCTTCCACCACGGGGAG
CCCAGCTGTCAGCCGCCTCACAGGAAGATGCTGCGTCGGCGGGGCAGCCCTGGCATGGGTGT
GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCCTGTGCTGCTCCTTCTCC
CCTGAGCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATAACAAACA
GCTGGTGCACAGCTTTGCTGAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCGCACGGCCC
TCTTCCCGGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCGTGCCTGTGGCG
GACGAGGGCAGCTTCACCTGCTTCGTGAGCATCCGGGATTTCGGCAGCGCTGCCGTGAGCCT
GCAGGTGGCCGCTCCCTACTCGAAGCCCAGCATGACCCTGGAGCCCAACAAGGACCTGCGGC
CAGGGGACACGGTGACCATCACGTGCTCCAGCTACCAGGGCTACCCTGAGGCTGAGGTGTTT
TGGCAGGATGGGCAGGGTGTGCCCCCTGACTGGCAACGTGACCACGTGCGAGATGGCCAACGA
GCAGGGCTTGTTTGATGTGCACAGCGTCCTGCGGGTGGTGCTGGGTGCGAATGGCACCTACA
GCTGCCTGGTGCGCAACCCCGTGCTGCAGCAGGATGCGCACRGCTCTGTCAACATCACAGGG
CAGCCTATGACATTCCCCCCAGAGGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCAT
TGCACTGCTGGTGGCCCTGGCTTTCGTGTGCTGGAGAAAGATCAAACAGAGCTGTGAGGAGG
AGAATGCAGGAGCTGAGGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT
CTGAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAATAGCCTGACCATGAGGACCAGG
GAGCTGCTACCCCTCCCTACAGCTCCTACCCTCTGGCTGCAATGGGGCTGCACTGTGAGCCC
TGCCCCCAACAGATGCATCCTGCTCTGACAGGTGGGCTCCTTCTCCAAAGGATGCGATACAC
AGACCACTGTGCAGCCTTATTTCTCCAATGGACATGATTCCCAAGTCATCCTGCTGCCTTTT
TTCTTATAGACACAATGAACAGACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCCTGCT
GCCTTATTTTACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATCACCACCTC
TTCTTCCAGTGCTGCGTGGACCATCTGGCTGCCTTTTTTCTCCAAAAGATGCAATATTCAGA
CTGACTGACCCCTGCCTTATTTACCAAAGACACGATGCATAGTACCCCGGCCCTTGTTTC
TCCAATGGCCGTGATACTAGTGATCATGTTTACGCCCTGCTTCCACCTGCATAGAATCTTT
TCTTCTCAGACAGGGACAGTGCGGCCTCAACATCTCCTGGAGTCTAGAAGCTGTTTCCTTTC
CCCTCCTTCCCTGCCCAAGTGAAGACAGGGCAGGGCCAGGAATGCTTTGGGGACACCG
AGGGGACTGCCCCCACCACCATGGTGCTATTCTGGGGCTGGGGCAGTCTTTTCTGCTGGC
TTGCCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCTTCCG
GATGTATCTCTCCCTGCCCCAGGAATGGAAGATGTGAGGACTTCTAATTTAAATGTGGGAC
TCGGAGGGATTTTGTAACCTGGGGGTATATTTTGGGGAAAATAAATGTCTTTGTAAAAAAA
AAAAAAAAAAAAA

FIGURE 51

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386

><subunit 1 of 1, 316 aa, 1 stop, 1 unknown

><MW: -1, pI: 4.62, NX(S/T): 4

MLRRRGSPGMGVHVGAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPPEPGFSLAQ
LNLIWQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQVRVADEGSFTCFV
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTITCSSYQGYPEAEVFWQDGQGVPL
TGNVTTSQMANEQGLFDVHSLRVVLGANGTYSCLVARNPVLQQDAHXSVTITGQPMTFPPEA
LWVTVGLSVCLIALLLVALAFVCWRKIKQSCEEENAGAEDQDGEGEGSKTALQPLKHSDSKED
DGQEIA

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

FIGURE 52

TTCGTGACCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCTTCTAAGAAGGGGGAGTC
CTGAACTTGTCTGAAGCCCTTGTCCGTAAGCCTTGAACCTACGTTCTTAAATCTATGAAGTCG
AGGGACCTTTTCGCTGCTTTTGTAGGGACTTCTTTCTTGCTTCAGCAACATGAGGCTTTTCT
TGTGGAACGCGGTCTTGACTCTGTTTCGTCACTTCTTTGATTGGGGCTTTGATCCCTGAACCA
GAAGTGAAAATTGAAGTTCTCCAGAAGCCATTCTGCCATCGCAAGACCAAAGGAGGGGA
TTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACTCCACTC
ACAAACATAACAATGGTCAGCCCATTTGGTTTACCCTGGGCATCCTGGAGGCTCTCAAAGGT
TGGGACCAGGGCTTGAAAGGAATGTGTGTAGGAGAGAAGAGAAAGCTCATCATTCCTCCTGC
TCTGGGCTATGGAAAAGAAGGAAAAGGTAAAATTCCTCCAGAAAGTACACTGATATTTAATA
TTGATCTCCTGGAGATTGCAAAATGGACCAAGATCCCATGAATCATTCCAAGAAATGGATCTT
AATGATGACTGGAACTCTCTAAAGATGAGGTTAAAGCATATTTAAAGAAGGAGTTTGAAAA
ACATGGTGCAGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTGTATAAAG
AAGATGAAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTA
TAGAGATACATCTACCTTTTAAATATAGCACTCATCTTCAAGAGAGGGCAGTCATCTTTAA
AGAACATTTTATTTTTATACAATGTTCTTTCTTGCTTTGTTTTTATTTTTATATATTTTTT
CTGACTCCTATTTAAAGAACCCTTAGGTTTCTAAGTACCCATTTCTTTCTGATAAGTTATT
GGGAAGAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTTCTACTTTTACAG
ATATGAAGCTTTGTTTTACTTTCTCACTTATAAATTTAAATGTTGCAACTGGGAATATACC
ACGACATGAGACCAGGTTATAGCACAAATTAGCACCTATATTTCTGCTTCCCTCTATTTTC
TCCAAGTTAGAGGTCAACATTTGAAAAGCCTTTTGCAATAGCCCAAGGCTTGCTATTTTCAT
GTTATAATGAAATAGTTTATGTGTAAGTGGCTCTGAGTCTCTGCTTGAGGACCAGAGGAAAA
TGGTTGTTGGACCTGACTTGTTAATGGCTACTGCTTTACTAAGGAGATGTGCAATGCTGAAG
TTAGAAACAAGGTTAATAGCCAGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAG
GCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTCGAGACCAGCCTGACCAACACGGAGAAA
CCCTATCTCTACTAAAAATACAAAGTAGCCCGGCGTGGTGATGCGTGCTGTAATCCCAGCT
ACCCAGGAAGGCTGAGGCGGCAGAATCACTTGAACCCGAGGCGAGGTTGCGGTAAGCCGAG
ATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAGAAAAGAACACGGTTAATACCATATNA
ATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGGCTCCTAGTGAT
TGGTGGCCTATTATGATAAATAGGACAAATCATTTATGTGTGAGTTTCTTTGTAATAAAATG
TATCAATATGTTATAGATGAGGTAGAAAGTTATATTTATATTCAATATTTACTTCTTAAGGC
TAGCGGAATATCCTTCCTGGTTCTTTAATGGGTAGTCTATAGTATATTATACTACAATAACA
TTGTATCATAAGATAAAGTAGTAAACCAGTCTACATTTTCCCATTTCTGTCTCATCAAAAAC
TGAAGTTAGCTGGGTGTGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGGGCCAAGGAGGG
TGGATCACTTGAGATCAGGAGTTCAGACACAGCCTGGCCAACATGGTGAAACCTTGTCTCTA
CTAAAAATACAAAATTAGCCAGGCGTGGTGGTGCACACCTGTAGTCCCAGCTACTCGGGAG
GCTGAGACAGGAGATTTGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCC
ACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAGAAGCAGA
CCTACAGCAGCTACTATTGAATAAATACCTATCCTGGATTTT

FIGURE 53

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44194

><subunit 1 of 1, 211 aa, 1 stop

><MW: 24172, pI: 5.99, NX(S/T): 1

MRLFLWNAVLTTLFVTSLIGALIPPEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGS
L
FHSTHKHNNGQPIWFTLGILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGI
P
PEST
LIFNIDLLLEIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKHGAVVNESH
H
DALVED
IFDKEDDKDGFISAREFTYKHDEL

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

FIGURE 54

AATAAAGCTTCCTTAATGTTGTATATGTCTTTGAAGTACATCCGTGCATTTTTTTTTTAGCAT
CCAACCATTCCTCCCTTGTAAGTTCTCGCCCCCTCAAATCACCTCTCCCGTAGCCCAACCGA
CTAACATCTCAGTCTCTGAAAATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCT
CACGGGGCTCAGTCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCACAGTAC
CTGCCACCCTCAACGTCTCAATGGCTCTGACGCCCGCCTGCCCTGCACCTTCAACTCCTGC
TACACAGTGAACCACAAACAGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTC
TGAGGAGATGTTCCCTCCAGTTCGCGATGAAGATCATTAACCTGAAGCTGGAGCGGTTTTCAAG
ACCGCGTGGAGTTCTCAGGGAACCCAGCAAGTACGATGTGTGCGGTGATGCTGAGAAACGTG
CAGCCGGAGGATGAGGGGATTTACAACCTGCTACATCATGAACCCCCCTGACCGCCACCGTGG
CCATGGCAAGATCCATCTGCAGGTCTCATGGAAGAGCCCCCTGAGCGGGACTCCACGGTGG
CCGTGATTGTGGGTGCCTCCGTGCGGGGCTTCTGGCTGTGGTTCATCTTGGTGCTGATGGTG
GTCAAGTGTGTGAGGAGAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGA
GGAGGGCAAGACGGACGGTGAAGGCAACCCGGATGATGGCGCCAAGTAGTGGGTGGCCGGCC
CTGCAGCCTCCCGTGTCCCGTCTCCTCCCCTCTCCGCCCTGTACAGTGACCTGCCTGCTCG
CTCTTGGTGTGCTTCCCGTGACCTAGGACCCAGGGCCACCTGGGGCCTCCTGAACCCCCG
ACTTCGTATCTCCACCCCTGCACCAAGAGTGACCCACTCTCTTCCATCCGAGAAACCTGCCA
TGCTCTGGGACGTGTGGGCCCTGGGGAGAGGAGAGAAAGGGCTCCCACCTGCCAGTCCCTGG
GGGGAGGCAGGAGGCACATGTGAGGGTCCCAGAGAGAAGGGAGTGGGTGGGCAGGGGTAGA
GGAGGGGCCGCTGTACCTGCCAGTGCTTGCCTGGCAGTGGCTTCAGAGAGGACCTGGTGG
GGAGGGAGGGCTTTCCTGTGCTGACAGCGCTCCCTCAGGAGGGCCTTGGCCTGGCACGGCTG
TGCTCCTCCCCTGCTCCCAGCCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTTCCTGA
AACTTGAGGGGCATGTTAAAGGGATGACTGTGCATTCCAGGGCACTGACGGAAGCCAGGG
CTGCAGGCAAAGCTGGACATGTGCCCTGGCCCAGGAGGCCATGTTGGGCCCTCGTTTCCATT
GCTAGTGGCCTCCTTGGGGCTCCTGTTGGCTCCTAATCCCTTAGGACTGTGGATGAGGCCAG
ACTGGAAGAGCAGCTCCAGGTAGGGGGCCATGTTTCCAGCGGGGACCCACCAACAGAGGCC
AGTTTCAAAGTCAGCTGAGGGGCTCAGGGGTGGGGCTCCATGGTGAATGCAGGTGTGCTGCAG
GCTCTGCCTTCTCCATGGGGTAACCACCCTCGCCTGGGCAGGGGCAGCCAAGGCTGGGAAAT
GAGGAGGCCATGCACAGGGTGGGGCAGCTTCTTTGGGGCTTCAGTGAGAACTCTCCCAGTT
GCCCTTGGTGGGGTTTCCACCTGGCTTTTGGCTACAGAGAGGGAAGGGAAAGCCTGAGGCCG
GCATAAGGGGAGGCCTTGGAACCTGAGCTGCCAATGCCAGCCCTGTCCCATCTGCGGCCACG
CTACTCGCTCCTCTCCCAACAACCTCCCTTCGTGGGGACAAAAGTGACAATTGTAGGCCAGGC
ACAGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGGATTACCTCCAT
CTGTTTAGTAGAAATGGGCAAAACCCCATCTCTACTAAAATAACAAGAATTAGCTGGGCGTG
GTGGCGTGTGCCTGTAATCCCAGCTATTGGGAGGCTGAGGCAGGAGAATCGCTTGAGCCCG
GGAAGCAGAGGTTGCAGTGAAGTGAAGTAGTATAGTGCCACTGCAATTCAGCCTGGGTGAC
ATAGAGAGACTCCATCTCAAAAAAA

FIGURE 55

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415

<subunit 1 of 1, 215 aa, 1 stop

<MW: 24326, pI: 6.32, NX(S/T): 4

MHRDAWLPRPAFSLTGSLFFSLVPPGRSMEVTVPATLNVLNGSDARLPCTFNSCYTVNHKQ
FSLNWTYQECNNCSEEMFLQFRMKIINLKLERFQDRVEFSGNPSKYDVSVMLRNVQPEDEGI
YNCYIMNPPDRHRHGKIHLLQVLMEEPPERDSTVAVIVGASVGGFLAVVILVLMVVKCVRRK
KEQKLSTDDLKTEEEGKTDGEGNPDDGAK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 161-179

Immunoglobulin-like fold:

amino acids 83-127

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

FIGURE 56

GTTGTATATGTCCTGAAGTACATCCGTGCATTTTTTTTTTAGCATCCAACCATCCTCCCTTGTA
GTTCTCGCCCCCTCAAATCACCTTCTCCCTTAGCCCACCCNACTAACATCTCAGTCTCTGAA
AATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCTCACGGGGCTCAGTCTCTTTT
TCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCCACAGTACCTGNCCACCCTCAACGTCC
TCAATGGCTCTGACGCCCCGCCTGCCCTGCCCTTCAACTCCTGCTACACAGTGAACCACAAAC
AGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTCTGAGGAGATGTTCTCCAG
TTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAGACCGCGTGGAGTTCTCAGG
GAACCCAGCAAGTACGATGTGTGCGGTGATGCTGAGAAACGTGCAGCCGGAGGATGAGGGGA
TTTACAAC TGCTACATCATGAACCCCCC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 57

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTNCACATA
CCTGCCACCCTCAACGTCCTCAATGGCTTTGACGCCCGCCTGCCCTGCACCTTCAACTCCNG
CTACACAGTGAACCACAAACAGTTCTCCCTGAACTGGATTACCAGGAGTGCAACAACCTGGC
TCTGAGGAGATGTTCTCCTCCAGTTCCCGCATGGAAGATCATTTAACCTGAAAGCTGGAAGCGG
TTTTCAAGAACCGCGTGGAAGTTTCTCAGGGAACCCAGCAAGTACGATGTGTCGGTGATGC
TGAGAAACGTGCAGCCGGAGGATGAGGGGATTTACAACTGCTACATCATGAACCCCCC

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTNCACATA

FIGURE 58

TGCGGCGACCGTCGTACACCAATGGGCTCCACCTCCGCCCCCTACCGTGTGGGGCTGCTCCCGGATGGCCTCCTGT
TCCTCTTGCTGCTGCTAATGCTGCTCGCGGACCCAGCGCTCCCGCCGGACGTCAACCCCAAGTGGTGCTGGTCC
CTGGTGATTTGGGTAACCAACTGGAAGCCAAGCTGGACAAGCCGACAGTGGTGCACTACCTCTGCTCCAAGAAGA
CCGAAAGCTACTTCACAATCTGGCTGAACCTGGAACCTGCTGCTGCTCATCATTGACTGCTGGATTGACAATA
TCAGGCTGGTTTACAACAAAACATCCAGGGCCACCCAGTTTCTGATGGTGTGGATGTACGTGTCCCTGGCTTTG
GGAAGACCTTCTCACTGGAGTTCTTGGACCCAGCAAAAGCAGCGTGGGTTCTATTTCCACACCATGGTGGAGA
GCCTTGTGGGCTGGGGCTACACACGGGGTGAGGATGTCCGAGGGGCTCCCTATGACTGGCGCCGAGCCCCAAATG
AAAACGGGCCCTACTTCTGGCCCTCCGCGAGATGATCGAGGAGATGTACCAGCTGTATGGGGGCCCCGTGGTGC
TGTTTGCCCAAGTATGGGCAACATGTACACGCTCTACTTTCTGCAGCGGCAGCCGAGCCGAGGCTGGAAGGACAAGT
ATATCCGGGCCCTTCGTGTCACTGGGTGCGCCCTGGGGGGCGTGGCCAAGACCTGCGCGTCTGGCTTCAGGAG
ACAACAACCGGATCCAGTCATCGGGCCCCCTGAAGATCCGGGAGCAGCAGCGGTGAGCTGTCTCCACAGCTGGC
TGCTGCCCTACAACATACATGGTCACCTGAGAAGGTGTTCTGTCAGACACCCACAATCAACTACACACTGCGGG
ACTACCGCAAGTTCTTCCAGGACATCGGCTTTGAAGATGGCTGGCTCATGCGGCAGGACACAGAAGGGCTGGTGG
AAGCCACGATGCCACCTGGCGTGCAGCTGCACTGCCCTCTATGGTACTGGCGTCCCAACACAGACTCCTTCTACT
ATGAGAGCTTCCCTGACCGTGACCTAAAATCTGCTTTGGTGACGGCGATGGTACTGTGAACCTGAAGAGTGCCC
TGCAGTGCCAGGCCTGGCAGAGCCGCGAGGAGACCAAGTGTGCTGTCAGGAGCTGCCAGGCAGCGAGCATCG
AGATGCTGGCCAACGCCACCACCTGGCCTATCTGAAACGTGTGCTCCTTGGGCCCCTGA

CTCCTGTGGCTCGGCCGTGGACCTGCTGTTGGCCTCTGGGGCTGTCTATGGCCACGCGTTTGGCAAAGTTTGTGA
CTCACCATTCAAGGCCCCGAGTCTTGGACTGTGAAGCATCTGCCATGGGGAAGTGCTGTTTGTATCCTTTCTCT
GTGGCAGTGAAGAAGGAAGAAATGAGAGTCTAGACTCAAGGGACACTGGATGGCAAGAATGCTGCTGATGGTGA
ACTGCTGTGACCTTAGGACTGGCTCCACAGGGTGGACTGGCTGGGCCCCGTGGTCCAGTCCCTGCCTGGGGCCATG
TGTCCCCCTATTCTGTGGGCTTTTCACTACTGGCTACTGGGCCCTGGCCCCGAGCCTTCTATGAGGGATGTT
ACTGGGCTGTGGTCTGTACCCAGAGGTCCAGGGATCGGCTCCTGGCCCCCTCGGGTGACCTTCCACACACCA
GCCACAGATAGGCCTGCCACTGGTTCATGGGTAGCTAGAGCTGCTGGCTTCCCTGTGGCTTAGCTGGTGGCCAGCC
TGACTGGCTTCTTGGGCGAGCCTAGTAGCTCCTGCAGGCAGGGGAGTTTGTTCGTTCTTCGTGGTTCCAGGC
CCTGGGACATCTCACTCCACTCCTACCTCCCTTACCACCAGGAGCATTCAGCTCTGGATGGGCAGCAGATGTG
CCCCAGTCCCGCAGGCTGTGTTCCAGGGGCCCTGATTTCTCGGATGTGCTATTGGCCCCAGGACTGAAGCTGC
CTCCCTTACCCTGGGACTGTGGTTCCAAGGATGAGAGCAGGGGTTGGAGCCATGGCCTTCTGGGAACCTATGGA
GAAAGGGAATCCAAGGAAGCAGCCAAGGCTGCTCGCAGCTTCCCTGAGCTGCACCTCTTGCTAACCCACCATCA
CACTGCCACCCTGCCCTAGGGTCTCACTAGTACCAAGTGGGTGAGCACAGGGCTGAGGATGGGGCTCCTATCCAC
CCTGGCCAGCACCCAGCTTAGTGCTGGGACTAGCCAGAACTTGAATGGGACCTGAGAGAGCCAGGGGTCCCC
TGAGGCCCCCTAGGGGCTTTCTGTCTGCCCCAGGGTGTCCATGGATCTCCCTGTGGCAGCAGGCATGGAGAGT
CAGGGCTGCCTTCATGGCAGTAGGCTCTAAGTGGGTGACTGGCCACAGGCCGAGAAAAGGGTACAGCCTCTAGGT
GGGGTTCCTCAAAGACGCCCTTCCAGGCTGGACTGAGCTGCTCTCCACAGGGTTTCTGTGACGCTGGATTTTCTCTG
TTGCATACATGCCTGGCATCTGTCTCCCTTGTTCCTGAGTGGCCCCACATGGGGCTCTGAGCAGGCTGTATCTG
GATTCGGCAATAAAAGTACTCTGGATGCTGTAAAAA

FIGURE 59

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189

><subunit 1 of 1, 412 aa, 1 stop

><MW: 46658, pI: 6.65, NX(S/T): 4

MGLHLRPYRVGLLPDGLLFLLLLLLMLLADPALPAGRHPVVLVPGDLGNQLEAKLDKPTVVH
YLCSKKTESYFTIWLNLLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDRVPFGKTFSL
EFLDPSKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEMYQ
LYGGPVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRI
PVIGPLKIREQQRSVSTSWLLPYNYTWSPEKVVFQTPTINYTLRDYRKFFQDIGFEDGWL
MQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQCQ
AWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLLGP

Important features:

Signal peptide:

amino acids 1-28

Potential lipid substrate binding site:

amino acids 147-164

N-glycosylation sites.

amino acids 99-102, 273-276, 289-292 and 398-401

Lipases, serine proteins

amino acids 189-201

Beta-transducin family Trp-Asp repeat

amino acids 353-365

FIGURE 60

CGGACGCGTGGGCGGACGCGTGGGGCGGCGGCAGCGGCGGCGACGGCGACATGGAGAGCGGG
GCCTACGGCGCGGCCAAGGCGGGCGGCTCCTTCGACCTGCGGCGCTTCCTGACGCAGCCGCA
GGTGGTGGCGCGCGCCGTGTGCTTGGTCTTCGCCTTGATCGTGTTCTCCTGCATCTATGGTG
AGGGCTACAGCAATGCCCACGAGTCTAAGCAGATGTACTGCGTGTTCAACCGCAACGAGGAT
GCCTGCCGCTATGGCAGTGCCATCGGGGTGCTGGCCTTCCTGGCCTCGGCCTTCTTCTTGGT
GGTCGACGCGTATTTCCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTCATTGGTG
ACCTGCTCTTCTCAGCTCTCTGGACCTTCCTGTGGTTTGTGGTTTCTGCTTCCTCACCAAC
CAGTGGGCAGTCACCAACCCGAAGGACGTGCTGGTGGGGGCCGACTCTGTGAGGGCAGCCAT
CACCTTCAGCTTCTTTTCCATCTTCTCCTGGGGTGTGCTGGCCTCCCTGGCCTACCAGCGCT
ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCACTCCGGACCCCAACACT
GCCTACGCCTCCTACCCAGGTGCATCTGTGGACAACCTACCAACAGCCACCCTTCACCCAGAA
CGCGGAGACCACCGAGGGCTACCAGCCGCCCCCTGTGTACTTGAGTGGCGGTTAGCGTGGGAA
GGGGGACAGAGAGGGCCCTCCCCCTCTGCCCTGGACTTTCCCATCAGCCTCCTGGAACCTGCCA
GCCCCCTCTTTTACCTGTTCCATCCTGTGCAGCTGACACACAGCTAAGGAGCCTCATAGCC
TGGCGGGGGCTGGCAGAGCCACACCCCAAGTGCCTGTGCCCAGAGGGCTTCAGTCAGCCGCT
CACTCCTCCAGGGCACTTTTAGGAAAGGGTTTTTAGCTAGTGTTTTTCTCGCTTTTAATGA
CCTCAGCCCCCGCCTGCAGTGGCTAGAAGCCAGCAGGTGCCCATGTGCTACTGACAAGTGCCT
CAGCTTCCCCCGGCCCGGGTCAGGCCGTGGGAGCCGCTATTATCTGCGTTCTCTGCCAAAG
ACTCGTGGGGGCCATCACACCTGCCCTGTGCAGCGGAGCCGGACCAGGCTCTTGTGTCCTCA
CTCAGGTTTGCTTCCCCTGTGCCCACTGCTGTATGATCTGGGGGCCACCACCCTGTGCCGGT
GGCCTCTGGGCTGCCTCCCGTGGTGTGAGGGCGGGGCTGGTGCTCATGGCACTTCCTCCTTG
CTCCCACCCCTGGCAGCAGGGAAGGGCTTTGCCTGACAACACCAGCTTTATGTAAATATTC
TGCAGTTGTTACTTAGGAAGCCTGGGGAGGGCAGGGGTGCCCCATGGCTCCAGACTCTGTC
TGTGCCGAGTGATTATAAAATCGTGGGGGAGATGCCCGGCCTGGGATGCTGTTTGGAGACG
GAATAAATGTTTTCTCATTCAAAG

FIGURE 61

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48304

<subunit 1 of 1, 224 aa, 1 stop

<MW: 24810, pI: 4.75, NX(S/T): 1

MESGAYGAAKAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGEYSNAHESKQMYCVFN
RNEDACRYGSAIGVLAFLASAFFLVVDAYFPQISNATDRKYLVIGDLLFSALWTFWLWVGFC
FLTNQWAVTNPKDVLVGADSVRAAITFSFFSIFSWGVLASLAYQRYKAGVDDFIQNYVDPTP
DPNTAYASYPGASVDNYQQPPFTQNAETTEGYQPPPVY

Important features:

Type II Transmembrane domain:

amino acids 1-45

Other transmembrane domains:

amino acids 74-90, 108-126 and 145-161

N-glycosylation site.

amino acids 97-100

FIGURE 62

GAGCCACCTACCCTGCTCCGAGGCCAGGCCTGCAGGGCCTCATCGGCCAGAGGGTGATCAGTGAGCAGAAGGATG
 CCCGTGGCCGAGGCCCCCCAGGTGGCTGGCGGGCAGGGGGACGGAGGTGATGGCGAGGAAGCGGAGCCAGAGGGG
 ATGTTCAAGGCCCTGTGAGGACTCCAAGAGAAAAGCCCCGGGGCTACCTCCGCCTGGTGCCCCCTGTTTGTGCTGCTG
 GCCCTGCTCGTGCTGGCTTCGGCGGGGGTGCTACTCTGGTATTTCTAGGGTACAAGGCGGAGGTGATGGTCAGC
 CAGGTGTACTCAGGCAGTCTGCGTGTACTCAATCGCCACTTCTCCAGGATCTTACCCGCCGGGAATCTAGTGCC
 TTCCGCAGTGAAACCGCCAAAGCCCAAGATGCTCAAGGAGCTCATCACCAGCACCCGCCCTGGGAACCTTACTAC
 AACTCCAGCTCCGTCTATTCTTTGGGGAGGGACCCCTCACCTGCTTCTTCTGGTTTATTCTCCAAATCCCCGAG
 CACCGCCGGCTGATGCTGAGCCCCGAGGTGGTGAGGCACTGCTGGTGGAGGAGCTGCTGTCCACAGTCAACAGC
 TCGGCTGCCGTCCCCCTACAGGGCCGAGTACGAAGTGGACCCCGAGGGCCTAGTGATCCTGGAAGCCAGTGTGAAA
 GACATAGCTGCATTGAATTCCACGCTGGGTGTGTACCGCTACAGCTACGTGGGCCAGGGCCAGGTCTCCGGCTG
 AAGGGGCCCTGACCACCTGGCCCTCCAGCTGCCTGTGGCACCTGCAGGGCCCCAAGGACCTCATGCTCAAACCTCCGG
 CTGGAGTGGACGCTGGCAGAGTGGCCGGACCGACTGGCCATGTATGACGTGGCCGGGGCCCCCTGGAGAAGAGGGCTC
 ATCAGCTCGGTGTACGGCTGCAGCCGCCAGGAGCCCGTGGTGGAGGTTCTGGCGTCCGGGGGCCATCATGGCGGT
 GTCTGGAAGAAGGGCCTGCACAGCTACTACGACCCCTTCGTGCTCTCCGTGCAGCCGGTGGTCTTCCAGGCCCTGT
 GAAGTGAACCTGACGCTGGACAACAGGCTCGACTCCAGGGCGTCTCAGCACCCCGTACTTCCCCAGCTACTCAG
 TCGCCCCAAACCCACTGCTCCTGGCACCTCACGGTGGCCCTCTCTGGACTACGGCTTGGCCCTCTGGTTTGTATGCC
 TATGCACTGAGGAGGCAGAAGTATGATTTGCCGTGCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGT
 GGCTTGGCATCCTGCAGCCCTACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACC
 TCCCAGATCTCCCTACCGGGCCCCGTGTGCGGTGCACCTATGGCTTGTACAACAGTCCGACCCCTGCCCTGGA
 GAGTTCTCTGTTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGGTCAAGGACTGCCCCAACGGCCTGGAT
 GAGAGAACTGCGTTTGCAGAGCCACATTCCAGTGCAAAGAGGACAGCACATGCATCTCACTGCCCAAGGTCTGT
 GATGGGCAGCCTGATTGTCTCAACGGCAGCGATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACACC
 TTCCAGTGTGAGGACCGGAGCTGCGTGAAGAAGCCCAACCCGAGTGTGATGGGCGGGCCGACTGCAGGGACGGC
 TCGGATGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCCCTCAGCCGCATTGTTGGTGGAGCTGTGTCTCCGAG
 GGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTCCGGGTGACACATCTGTGGGGGGGCCCTCATCGCTGACCGC
 TGGGTGATAACAGCTGCCCCACTGCTTCCAGGAGGACAGCATGGCCTCCACGGTGTGTGGACCGTGTCTTGGGC
 AAGGTGTGGCAGAACTCGCGCTGGCCTGGAGAGGTGTCTTCAAGGTGAGCCCGCTGCTCCTGCACCCGTACAC
 GAAGAGGACAGCCATGACTACGACGTGGCGCTGCTGCAGCTCGACACCCGGTGGTGCCTCGGCCGCCGTGCGC
 CCCGTCTGCCTGCCCCGCGCTCCCACTTCTTCGAGCCCGCCTGCACCTGCTGGATTACGGGTGGGGCGCCTTG
 CGCGAGGGCGGCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTGATCCCAAGGACCTGTGCAGCGAG
 GCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGCCGCTACCGCAAGGGCAAGAAGGATGCCTGTAGGGT
 GACTCAGGTGGTCCGCTGGTGTGCAAGGCACTCAGTGGCCGCTGGTTCTTGGCGGGGCTGGTCACTGGGGCCTG
 GGCTGTGGCCGGCCTAACTACTTTCGGGCTCTACACCCGATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGTG
 ACCTGAGGAAGTGGCCCCCTGCAAAGCAGGGCCCCACCTCTGACTCAGAGAGCCAGGGCAACTGCCAAGCAGG
 GGGACAAGTATTCTGGCGGGGGGTGGGGGAGAGAGCAGGCCCTGTGGTGGCAGGAGGTGGCATCTTGTCTCGTCC
 CTGATGTCTGCTCCAGTGATGGCAGGAGGATGGAGAAGTGCCAGCAGCTGGGGGTCAAGACGTCCCCTGAGGACC
 CAGGCCCCACCCAGCCCTTCTGCCTCCCAATTCTCTCTCTCCGTCCCCCTTCTTCCACTGCTGCCTAATGCAAG
 GCAGTGGCTCAGCAGCAAGAATGCTGGTTCTACATCCCGAGGAGTGTCTGAGGTGCGCCCCACTCTGTACAGAGG
 CTGTTTGGGCAGCCTTGCCTCCAGAGAGCAGATTCCAGCTTCGGAAGCCCCCTGGTCTAACTTGGGATCTGGGAAT
 GGAAGGTGCTCCCATCGGAGGGGACCTCAGAGCCCTGGAGACTGCCAGGTGGGCCTGCTGCCACTGTAAGCCAA
 AAGGTGGGGAAGTCTGACTCCAGGTCCTTGGCCCCACCCCTGCCTGCCACCTGGGCCCTCACAGCCAGACCCCT
 CACTGGGAGGTGAGCTCAGCTGCCCTTTGGAATAAAGCTGCCTGATCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 63

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152

><subunit 1 of 1, 802 aa, 1 stop

><MW: 88846, pI: 6.41, NX(S/T): 7

MPVAEAPQVAGGQGDGGDGEEAEPEGMFKACEDSKRKARGYLRLVPLFVLLALLVLASAGVL
LWYFLGYKAEVMVSQVYSGSLRVLNRHFSQDLTRRESSAFRSETAKAQKMLKELITSTR LGT
YNNSSSVYSFGEGPLTCFFWFILQIPEHRRMLLSPEVVQALLVEELLSTVNSSAAVPYRAEY
EVDPEGLVILEASVKDIAALNSTLGCYRYSYVGQGQVLRRLKGPDH LASSCLWHLQGP KDLML
KLRLEWTLAECDRLAMYDVAGPLEKRLITSVYGCSRQEPVVEVLASGAIMAVVWKKGLHSY
YDPFVLSVQPVVFQACEVNLTLDNRLDSQGV LSTPYFPSYSPQTHCSWHLTVPSLDYGLAL
WFDAYALRRQKYDLPCTQGQWTIQNRRLCGLRILQPYAERIPVVATAGITINFTSQISLTGP
GVRVHYGLYNQSDPCPGEFLCSVNGLCVPACDGVKDCPNGLDERNCVCRATFQCKEDSTCIS
LPKVCDDGQPDCLNGSDEEQCQEGVPCGTFTFQCEDRSCVKKPNPQCDGRPD CRDGSDEEHCD
CGLQGPSSRIVGGAVSSEGEWPWQASLQVRGRHICGGALIADRWVITA AHCFQEDSMAS TVL
WTVFLGK VWQNSRWPGEVSFKVSRLLLHPYHEEDSHDYDVALLQLDHPVVRSA AAVRPVCLPA
RSHFFEPGLHCWITGWGALREGGPISNALQKVDVQLIPQDLCSEAYRYQVTPRMLCAGYRKG
KKDACQGDSSGGPLVCKALSGRWFLAGLVSWGLGCGRPNYFGVYTRITGVISWIQQVVT

Important features:

Type II transmembrane domain:

amino acids 46-67

Serine proteases, trypsin family, histidine active site.

amino acids 604-609

N-glycosylation sites.

amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447
and 509-512

Kringle domains.

amino acids 746-758 and 592-609

Homologous region to Kallikrein Light Chain:

amino acids 568-779

Homologous region to Low-density lipoprotein receptor:

amino acids 451-567

FIGURE 64

GCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGTGGCTTGCGCATCCTGCAGCCC
TACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACCTCCCAGAT
CTCCCTCACCGGGCCCCGGTGTGCGGGTGCACCTATGGCTTGTACAACCAGTCGGACCCCTGCC
CTGGAGAGTTCTCTGTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGGTCAAGGAC
TGCCCCAACGGCCTGGATGAGAGAACTGCGTTTGCAGAGCCACATTCCAGTGCAAAGAGGA
CAGCACATGCATCTCACTGCCCAAGGTCTGTGATGGGCAGCCTGATTGTCTCAACGGCAGCG
ATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACCTTCCAGTGTGAGGACCGG
AGCTGCGTGAAGAAGCCCAACCCGCAGTGTGATGGGCGGCCCGACTGCAGGGACGGCTCGGA
TGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCCCTCCAGCCGCATTGTTGGTGGAGCTGTGT
CCTCCGAGGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTCTGGGGTTCGACACATCTGTGGG
GGGGCCCTCATCGCTGACCGCTGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCAT
GGCCTCCACGGTGCTGTGGACCGTGTTCCTGGGCAAGGTGTGGCAGAACTCGCGCTGGCCTG
GAGAGGTGTCCTTCAAGGTGAGCCGCCTGCTCCTGCACCCGTACCAAGAGGACAGCCAT
GACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTTCGCTCGGCCCGCTGCGCCC
CGTCTGCCTGCCCCGCGCTCCCACTTCTTCGAGCCCGGCCTGCACTGCTGGATTACGGGCT
GGGGCGCCTTGCGCGAGGGCGGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTG
ATCCACAGGACCTGTGCAGCGAGGCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGC
CGGCTACCGCAAGGGCAAGAAGGATGCCTGTGAGGGTGACTCAGGTGGTCCGCTGGTGTGCA
AGGCACTCAGTGGCCGCTGGTTCTGCGGGGCTGGTCAGCTGGGGCCTGGGCTGTGGCCGG
CCTAACTACTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
GACCTGAGGAACTGCCCCCTGCAAAGCAGGGGCCACCTCCTGGACTCAGAGAGCCCAGGGC
AACTGCCAAGCAGGGGGACAAGTAT

FIGURE 65

GGACGAGGGCAGATCTCGTTCTGGGGCAAGCCGTTGACACTCGCTCCCTGCCACCGCCCGG
CTCCGTGCCGCCAAGTTTTCATTTTCCACCTTCTCTGCCTCCAGTCCCCCAGCCCCTGGCCG
AGAGAAGGGTCTTACCGGCCGGGATTGCTGGAAACACCAAGAGGTGGTTTTTGTTTTTTAAA
ACTTCTGTTTCTTGGGAGGGGGTGTGGCGGGGCAGGATGAGCAACTCCGTTCCCTCTGCTCTG
TTTCTGGAGCCTCTGCTATTGCTTTGCTGCGGGGAGCCCCGTACCTTTTGGTCCAGAGGGAC
GGCTGGAAGATAAGCTCCACAAACCCAAAGCTACACAGACTGAGGTCAAACCATCTGTGAGG
TTTAACCTCCGCACCTCCAAGGACCCAGAGCATGAAGGATGCTACCTCTCCGTGCGCCACAG
CCAGCCCTTAGAAGACTGCAGTTTCAACATGACAGCTAAAACCTTTTTTCATCATTACGGAT
GGACGATGAGCGGTATCTTTGAAAACCTGGCTGCACAACTCGTGTGAGCCCTGCACACAAGA
GAGAAAGACGCCAATGTAGTTGTGGTTGACTGGCTCCCCCTGGCCCACCAGCTTTACACGGA
TGCGGTCAATAATACCAGGGTGGTGGGACACAGCATTGCCAGGATGCTCGACTGGCTGCAGG
AGAAGGACGATTTTTCTCTCGGGAATGTCCACTTGATCGGCTACAGCCTCGGAGCGCACGTG
GCCGGGTATGCAGGCAACTTCGTGAAAGGAACGGTGGGCCGAATCACAGGTTTGGATCCTGC
CGGGCCCATGTTTGAAGGGGCCGACATCCACAAGAGGCTCTCTCCGGACGATGCAGATTTTG
TGGATGTCCTCCACACCTACACGCGTTCCTTCGGCTTGAGCATTGGTATTCAGATGCCTGTG
GGCCACATTGACATCTACCCCAATGGGGGTGACTTCCAGCCAGGCTGTGGACTCAACGATGT
CTTGGGATCAATTGCATATGGAACAATCACAGAGGTGGTAAAATGTGAGCATGAGCGAGCCG
TCCACCTCTTTGTTGACTCTCTGGTGAATCAGGACAAGCCGAGTTTTGCCTTCCAGTGCACT
GACTCCAATCGCTTCAAAAAGGGGATCTGTCTGAGCTGCCGCAAGAACCGTTGTAATAGCAT
TGGCTACAATGCCAAGAAAATGAGGAACAAGAGGAACAGCAAAATGTACCTAAAAACCCGGG
CAGGCATGCCTTTCAGAGGTAACCTTCAGTCCCTGGAGTGTCCCTGAGGAAGGCCCTTAATA
CCTCCTTCTTAATACCATGCTGCAGAGCAGGGCACATCCTAGCCCAGGAGAAGTGGCCAGCA
CAATCCAATCAAATCGTTGCAAATCAGATTACACTGTGCATGTCCTAGGAAAGGGAATCTTT
ACAAAATAAACAGTGTGGACCCCTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646

><subunit 1 of 1, 354 aa, 1 stop

><MW: 39362, pI: 8.35, NX(S/T): 2

MSNSVPLLCFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRRTSKDPEHE
GCYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVVDWL
PLAHQLYTDVANNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTV
GRITGLDPAGPMFEGADIAHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDF
QPGCGLNDVLGSIAYGTITEVVKCEHERAVHLFVDSLVDKPSFAFQCTDSNRFKKGICLS
CRKNRCNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

Important features:

Signal peptide:

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

FIGURE 67

CGGACGCGTGGGCGGACGCGTGGGCTGGGCAAGGGCCGGGGCGCGGGCCGAGCCACCTCTTCCCCCTCCCCCG
TTCCCTGTGCGCGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCACCCGGCCGGCCCTGGGGGGCTGACAGT
CGGCAAAGTTTGGCCCCGAAGAGGAAGTGGTCTCAAACCCCGGCAGGTGGCGACCAGGCCAGACCAGGGGCGCTCG
CTGCCTGCGGGCGGGCTGTAGGCGAGGGCGCGCCCCAGTGCCGAGACCCGGGGCTTCAGGAGCCGGCCCCGGGAG
AGAAGAGTGCGGCGGCGGACGGAGAAAACAACCTCAAAGTTGGCGAAAGGCACCGCCCCCTACTCCCGGGCTGCCG
CCGCCTCCCCGCCCCAGCCCTGGCATCCAGAGTACGGGTGAGACCCGGGCCATGGAGCCCCCTGGGGAGGCGG
CACCAGGGAGCCTGGGCGCCCCGGGGCTCCGCCGCGACCCCATCGGGTAGACCACAGAAGCTCCGGGACCCCTCCG
GCACCTCTGGACAGCCAGGATGCTGTTGGCCACCTCTCTCTCTCTCTCTTGGAGGCGCTCTGGCCCATCCAG
ACCGGATTATTTTTTCAAATCATGCTTGTGAGGACCCCCCAGCAGTGCTCTTAGAAGTGCAGGGCACCTTACAGA
GGCCCCCTGGTCCGGGACAGCCGCACCTCCCTTGCCAACTGCACCTGGCTCATCTGGGAGCAAGGAACAGACTG
TCACCATCAGGTTCCAGAAGCTACACCTGGCCTGTGGCTCAGAGCGCTTAACCTACGCTCCCCTCTCCAGCCAC
TGATCTCCCTGTGTGAGGCACCTCCCAGCCCTCTGCAGCTGCCCCGGGGCAACGTACCATCATTACAGCTATG
CTGGGGCCAGAGCACCCATGGGCCAGGGCTTCCTGCTCTCTACAGCCAAGATTGGCTGATGTGCC'TGCAGGAAG
AGTTTTCAGTGCC'TGAACCACCGCTGTGTATCTGCTGTCCAGCGCTGTGATGGGGTTGATGCC'TGTGGCGATGGCT
CTGATGAAGCAGGTTGCAGCTCAGACCCCTTCCCTGGCTGACCCCCAAGACCCGTCCCCTCCCTGCC'TTGCAATG
TCACCTTGGAGGACTTCTATGGGGTCTTCTCTCTCTCTGGATATACACACCTAGCCTCAGTCTCCACCCCCAGT
CCTGCCATTGGCTGCTGGACCCCCATGATGGCCGGCGGCTGGCCGTGCGCTTACAGCCCTGGACTTGGGCTTTG
GAGATGCAGTGCATGTGTATGACGGCCCTGGGCCCCCTGAGAGCTCCCGACTACTGCGTAGTCTCACCCACTTCA
GCAATGGCAAGGCTGTCACTGTGGAGACACTGTCTGGCCAGGCTGTTGTGTCTTACCACACAGTTGCTTGGAGCA
ATGGTCGTGGCTTCAATGCCACCTACCATGTGCGGGGCTATTGCTTGCCTTGGGACAGACCCCTGTGGCTTAGGCT
CTGGCCTGGGAGCTGGCGAAGGCTAGGTGAGCGCTGCTACAGTGAGGCACAGCGCTGTGACGGCTCATGGGACT
GTGCTGACGGCACAGATGAGGAGGACTGCCAGGCTGCCACCTGGACACTTCCCCTGTGGGGCTGCTGGCACCT
CTGGTGCCACAGCCTGCTACCTGCCCTGCTGACCGCTGCAACTACCAGACTTTCTGTGCTGATGGAGCAGATGAGA
GACGCTGTGCGCATTGGCAGCCTGGCAATTTCCGATGCCGGGACGAGAAGTGCCTGTATGAGACGTGGGTGTGCG
ATGGGCAGCCAGACTGTGCGGACGGCAGTGTGATGGGACTGCTCCTATGTTCTGCCCCGCAAGGTCAATTACAG
CTGCAGTCATTGGCAGCCTAGTGTGCGGCTGCTCCTGGTCACTCGCCCTGGGCTGCACCTGCAAGCTCTATGCCA
TTGCAACCCAGGAGTACAGCATCTTTGCCCCCTCTCCCGGATGGAGGCTGAGATTGTGCAGCAGCAGGCACCCC
CTTCTACGGGCAGCTCAATTGCCCAGGGTGCCATCCCACCTGTAGAAGACTTTCTACAGAGAATCCTAATGATA
ACTCAGTGCTGGGCAACCTGCGTCTCTGCTACAGATCTTACGCCAGGATATGACTCCAGGAGGTGGCCAGGTG
CCCGCCGTGCTCAGCGGGGCGCTTGATGCGACGCTGGTACGCCGTCTCCGCCGTGGGGCTTGCTCCCTCGAA
CCAACACCCCGGCTCGGGCCTCTGAGGCCAGATCCCAGGTACACCTTCTGCTGCTCCCCTTGAGGCCCTAGATG
GTGGCACAGGTCCAGCCGTGAGGGCGGGCAGTGGGTGGGCAAGATGGGGAGCAGGCACCCCACTGCCCATCA
AGGCTCCCCTCCCATCTGCTAGCACGTCTCCAGCCCCCACTACTGTCCCTGAAGCCCCAGGGCCACTGCCCTCAC
TGCCCCCTAGAGCCATCACTA'TGTCTGGAGTGGTGAGGCCCTGCGAGGCCGCTGTTGCCAGCCTGGGGCCCC
CAGGACCAACCCGGAGCCCCCTGGACCCACACAGCAGTCTTGGCCCTGGAAGATGAGGACGATGTGCTACTGG
TGCCACTGGCTGAGCCGGGGGTGTGGGTAGCTGAGGCAGAGGATGAGCCACTGCTTACCTGAGGGGACCTGGGGG
CTCTACTGAGGCCTCTCCCCTGGGGGCTCTACTCATAGTGGCACAACTTTTAGAGGTGGGTGAGCCTCCCCTCC
ACCACTTCTTCCCTGTCCCTGGATTTAGGGACTTGGTGGGCCTCCCGTTGACCCTATGTAGCTGCTATAAAGT
TAAGTGTCCCTCAGGCAGGGAGAGGGCTCACAGAGTCTCTCTGTACGTGGCCATGGCCAGACACCCAGTCCCT
TCACCACCACCTGCTCCCCACGCCACCACATTTGGGTGGCTGT'TTTTAAAAAGTAAAGTTCTTAGAGGATCATA
GGTCTGGACACTCCATCCTTGCCAAACCTCTACCCAAAAGTGGCCTTAAGCACCGGAATGCCAATTAAGTAGAGA
CCCTCCAGCCCCCAAGGGGAGGATTTGGGCAGAACCTGAGGT'TTTGCCATCCACAATCCCTCCTACAGGGCCTGG
CTCAGAAAAGAGTGCAACAAATGCTTCTATTCCATAGCTACGGCATTGCTCAGTAAGTTGAGGTCAAAAATAAA
GGAATCATACATCTC

FIGURE 68

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631

<subunit 1 of 1, 713 aa, 1 stop

<MW: 76193, pI: 5.42, NX(S/T): 4

MLLATLLLLLLGGALAHPDRIIFPNHACEDPPAVLLEVQGTLQRPLVRDSRTSPANCTWLIL
GSKEQTVTIRFQKLHLACGSERLTLRSPLQPLISLCEAPPSPLQLPGGNVTITYSYAGARAP
MGQGFLLSYSQDWLMCLQEEFQCLNHRCVSAVQRC DGVDACGDGSDEAGCSSDPFPGLT PRP
VPSLPCNVLTLED FYGVFSSPGYTHLASVSH PQSCHWLLDPHDGRR LAVRFTALDLGFGDAVH
VYDGP GPPESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGFNATYHVRGYCLP
WDRPCGLGSGLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPPGHFPCGAAGTSGAT
ACYLPADRCNYQTF CADGADERRCRHCQPGNFR CRDEKCVYETWVCDGQPCADGSDEWDCS
YVLPRKVITA AAVIGSLVCGLLL VIALGCTCKLYAIR TQEYSIFAPLSRMEAEIVQQQAPPSY
GQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGGPGARRRQGR LMRRLVR
RLRRWGLLPRTNTPARASEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPPLPIKA
PLPSASTSPAPTTVPEAPGPLPSLPLEPSLLSGVVQALRGRL LPSLGPPGPTRSPPGPHTAV
LALEDEDDVLLVPLAEPGVWVAEAEDEPLLT

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 69

CGAGCTGGGCGAGAAGTAGGGGAGGGCGGTGCTCCGCCGCGGTGGCGGTTGCTATCGCTTCG
CAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCA
GACGCGATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCCTTCTGCTTCAGTGTGAAAGG
CCACGTGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTTTATCATCGCAC
AAGCCCCTGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTTCATA
CTTTTATATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCTTTGCTTGATAT
TATCAACTCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAG
AAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATGCTGTCTT
GCCGACGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCAGCGGTCCTTACCAGAAAAA
GCCTGTGCATGAAAAAAAAGAAGTTTTGTAATTTTATATTACTTTTTTAGTTTGATACTAAGT
ATTAAACATATTTCTGTATTCTTCCAAAAAAAAAAAAAAAAAAAA

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645

><subunit 1 of 1, 152 aa, 1 stop

><MW: 17170, pI: 9.62, NX(S/T): 1

MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEPYIVITGFEVTVILFFILL
YVLRLDRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTTLTVGGGVFALVTAVCCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL

Important features:

Potential type II transmembrane domain:

amino acids 26-42

Other potential transmembrane domain:

amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern

amino acids 78-99 and 85-106

N-myristoylation site.

amino acids 110-115

Ribonucleotide reductase large subunit protein

amino acids 116-127

FIGURE 71

GGGCGAGAAGTAGGGGAGGGCGTGTTCCGCCGCGGTGGCGGTTGCTATCGTTTTGCAGAACC
TACTCAGGCAGCCAGNTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCAGACGCGA
TGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGTGAAAGGCCACGTG
AAGATGCTGCGGCTGGCACTAACTGNGACATCTATGACCTTTTTTTATNATCGCACAAGCCCC
TGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTTCATACTTTTAT
ATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCTTTGCTTGATATTATCAAC
TCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACCAC
AACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATGCTGTNTTGCCGAC

FIGURE 72

CAGCCCCGCGCGCCGGCCGAGTCGCTGAGCCGCGGCTGCCGGACGGGACGGGACCGGCTAGG
CTGGGCGCGCCCCCGGGCCCCGCGTGGGCATGGGCGCACTGGCCCCGGGCGCTGCTGCTGC
CTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCGCGCCCTTCACGCTG
CCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCC
TGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGG
GCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTAC
CTGGAGATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAG
TAACTTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGT
CTAGCACATACCGCTCCAAGGGCTTTGACGTACAGTGAAGTACACACAAGGAAGCTGGACG
GGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAA
CATTGCCACTATTTTTGAATCAGAGAATTTCTTTTGCCTGGGATTAAATGGAATGGAATAC
TTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC
CTGGTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC
CGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGT
ATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTG
AAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGC
CATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGG
AAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAG
CTGGCGTGCTGGACGAATTTCGAAACACCTTGGTCTTACTTCCCTAAATCTCCATCTACCT
GAGAGACGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTACGC
CCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTTCGGCATTTCCTCCATCCACAAAT
GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAA
GAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTT
CCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAG
CCCATTTTGTGGATTGTGTCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTT
AATCGTCCTGCTGCTGCTGCCGTTCCGGTGTGAGCGTCGCCCCCGTGACCCTGAGGTCGTCA
ATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAATGAATAGCCAGGCCTGACCTCAAGCAA
CCATGAACTCAGCTATTAAGAAAATCACATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCG
CTTTCTCCTGTGCCACCCGTCTTCAATCTCTGTTCTGCTCCAGATGCCTTCTAGATTAC
TGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAA
AAAAAAACTTCATTCTAA

FIGURE 73

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493

><subunit 1 of 1, 518 aa, 1 stop

><MW: 56180, pI: 5.08, NX(S/T): 2

MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLAL
ALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS
YIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIFESNF
FLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGSGTNNGS
LVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKAIVDSGTTLLR
LPQKVFDAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISIIYLRDENSRSFR
ITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA
EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLIVLLLLPFRCL
QRRPRDPEVVNDESSLVRHRWK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCCTTCGGAGGCTGACGCGCCCGGGCGCCGTTCCAGGCCTGTGCAGGGCGGATCG
GCAGCCGCCTGGCGGCGATCCAGGGCGGTGCGGGGCTGGGCGGGAGCCGGGAGGCGCGGCC
GGCATGGAGGCGCTGCTGCTGGGCGCGGGGTTGCTGCTGGGCGCTTACGTGCTTGTCTACTA
CAACCTGGTGAAGGCCCCGCGTGCGGCGGCATGGGCAACCTGCGGGGCGGCACGGCCGTGG
TCACGGGCGCCAACAGCGGCATCGGAAAGATGACGGCGCTGGAGCTGGCGCGCCGGGGAGCG
CGCGTGGTGCTGGCCTGCCGCAGCCAGGAGCGCGGGGAGGCGGCTGCCTTCGACCTCCGCCA
GGAGAGTGGGAACAATGAGGTTCATCTTCATGGCCTTGGAATTGGCCAGTCTGGCCTCGGTGC
GGGCCTTTGCCACTGCCTTTCTGAGCTCTGAGCCACGGTTGGACATCCTCATCCACAATGCC
GGTATCAGTTCCTGTGGCCGGACCCGTGAGGCGTTTAACCTGCTGCTTCGGGTGAACCATAT
CGGTCCCTTTCTGCTGACACATCTGCTGCTGCCTTGCCTGAAGGCATGTGCCCTTAGCCGCG
TGGTGGTGGTAGCCTCAGCTGCCCACTGTCGGGGACGTCTTGACTTCAAACGCCTGGACCGC
CCAGTGGTGGGCTGGCGGCAGGAGCTGCGGGCATATGCTGACACTAAGCTGGCTAATGTACT
GTTTGCCCGGGAGCTCGCCAACCAGCTTGAGGCCACTGGCGTCACCTGCTATGCAGCCCACC
CAGGGCCTGTGAACTCGGAGCTGTTCTGCGCCATGTTCTTGATGGCTGCGCCCACTTTTG
CGCCATTGGCTTGGCTGGTGCTCCGGGCACCAAGAGGGGGTGCCAGACACCCCTGTATTG
TGCTCTACAAGAGGGCATCGAGCCCCCTCAGTGGGAGATATTTTGCCAACCTGCCATGTGGAAG
AGGTGCCTCCAGCTGCCCCGAGACGACCGGGCAGCCCATCGGCTATGGGAGGCCAGCAAGAGG
CTGGCAGGGCTTGGGCCTGGGGAGGATGCTGAACCCGATGAAGACCCCCAGTCTGAGGACTC
AGAGGCCCCATCTTCTCTAAGCACCCCCCACCCTGAGGAGCCCACAGTTTTCTAACCTTACC
CCAGCCCTCAGAGCTCACCAGATTTGTCTAAGATGACGCACCGAATTGAGGCTAAAGTTGAG
CCTGAGATCCAGCTCTCCTAACCCTCAGGCCAGGATGCTTGCCATGGCACTTCATGGTCCTT
GAAAACCTCGGATGTGTGTGAGGCCATGCCCTGGACACTGACGGGTTTGTGATCTTGACCTC
CGTGGTTACTTTCTGGGGCCCCAAGCTGTGCCCTGGACATCTCTTTTCTGGTTGAAGGAAT
AATGGGTGATTATTTCTTCTGAGAGTGACAGTAACCCAGATGGAGAGATAGGGGTATGCT
AGACACTGTGCTTCTCGGAAATTTGGATGTAGTATTTTCAGGCCCCACCCTTATTGATTCTG
ATCAGCTCTGGAGCAGAGGCAGGGAGTTTGCAATGTGATGCACTGCCAACATTGAGAATTAG
TGAAC TGATCCCTTTGCAACCGTCTAGCTAGGTAGTTAAATTACCCCCATGTTAATGAAGCG
GAATTAGGCTCCCGAGCTAAGGGACTCGCCTAGGGTCTCACAGTGAGTAGGAGGAGGGCCTG
GGATCTGAACCCAAGGTCTGAGGCCAGGGCCGACTGCCGTAAGATGGGTGCTGAGAAGTGA
GTCAGGGCAGGGCAGCTGGTATCGAGGTGCCCCATGGGAGTAAGGGGACGCCTTCCGGGCGG
ATGCAGGGCTGGGGTCATCTGTATCTGAAGCCCCTCGGAATAAAGCGCGTTGACCGCCAAAA
AAAAAAAAAAAAAAAAAAAA

FIGURE 75

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227

<subunit 1 of 1, 377 aa, 1 stop

<MW: 40849, pI: 7.98, NX(S/T): 0

MEALLLGAGLLLGAYVLVYYNLVKAPPCGGMGNLRGRTAVVTGANSIGKMTALELARRGAR
VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLSSEPRLDILIHNA
ISSCGRTREAFNLLLRVNHIGPFLLTHLLLPCLKACAPSRVVVVASAAHCRGRLLDFKRLDRP
VVGWRQELRAYADTKLANVLFARELANQLEATGVTCYAAHPGPVNSELFRLHVPGWLRPLLR
PLAWLVLRAPRGGAQTPLYCALQEGIEPLSGRYFANCHVEEVPPAARDRAHRLWEASKRL
AGLGPGEDAEPDEDPQSEDSEAPSSLSTPHPEEPTVSQPYPSPQSSPDLSKMTHRIQAKVEP
EIQLS

Important features:

Signal peptide:

amino acids 1-16

Glycosaminoglycan attachment site.

amino acids 46-49

Short-chain alcohol dehydrogenase family

amino acids 37-49 and 114-124

FIGURE 76

GGAGGAGACAGCCTCCTGGGGGGCAGGGGTTCCCTGCCTCTGCTGCTCCTGCTCATCATGGGAGGCATGGCTCAG
GACTCCCCGCCCCAGATCCTAGTCCACCCCAAGGACAGCTGTTCCAGGGCCCTGGCCCTGCCAGGATGAGCTGC
CAAGCCTCAGGCCAGCCACCTCCACCATCCGCTGGTTGCTGAATGGGCAGCCCTGAGCATGGTGCCCCAGAC
CCACACCACCTCCTGCCTGATGGGACCCTTCTGCTGCTACAGCCCCCTGCCCGGGACATGCCACGATGGCCAG
GCCCTGTCCACAGACCTGGGTGTCTACACATGTGAGGCCAGCAACCGGCTTGGCACGGCAGTCAGCAGAGGCGCT
CGGCTGTCTGTGGCTGTCTCCGGGAGGATTTCCAGATCCAGCCTCGGGACATGGTGGCTGTGGTGGGTGAGCAG
TTTACTCTGGAATGTGGGCCGCCCTGGGGCCACCCAGAGCCCACAGTCTCATGGTGGAAAGATGGGAAACCCCTG
GCCCTCCAGCCCGGAAGGCACACAGTGTCCGGGGGGTCCCTGCTGATGGCAAGAGCAGAGAAGAGTGACGAAGGG
ACCTACATGTGTGTGGCCACCAACAGCGCAGGACATAGGGAGAGCCGCGCAGCCCGGGTTTCCATCCAGGAGCCC
CAGGACTACACGGAGCCTGTGGAGCTTCTGGCTGTGCGAATTTCAGCTGGAAAAATGTGACACTGCTGAACCCGGAT
CCTGCAGAGGGCCCCAAGCCTAGACCGGCGGTGTGGCTCAGCTGGAAGGTGAGTGGCCCTGCTGCGCCTGCCAA
TCTTACACGGCCTTGTTCAGGACCCAGACTGCCCGGGAGGCCAGGGAGCTCCGTGGGCAGAGGAGCTGCTGGCC
GGCTGGCAGAGCGCAGAGCTTGGAGGCCTCCACTGGGGCCAAAGACTACGAGTTCAAAGTGAGACCATCCTCTGGC
CGGGCTCGAGGCCCTGACAGCAACGTGCTGCTCCTGAGGCTGCCGAAAAAGTGCCAGTGCCCCACCTCAGGAA
GTGACTCTAAAGCCTGGCAATGGCACTGTCTTTGTGAGCTGGTCCCACCACCTGCTGAAAAACCAATGGCATC
ATCCGTGGCTACAGGTCTGGAGCCTGGGCAACACATCACTGCCACCAGCCAAGTGGACTGTAGTTGGTGAGCAG
ACCCAGCTGGAAATCGCCACCCATATGCCAGGCTCCTACTGCGTGCAAGTGGCTGCAGTCACTGGTGTGGAGCT
GGGGAGCCCAGTAGACCTGTCTGCCTCCTTTTAGAGCAGGCCATGGAGCGAGCCACCCAAGAACCAGTGAGCAT
GGTCCCTGGACCCTGGAGCAGCTGAGGGCTACCTTGAAGCGGCTGAGGTCAATTGCCACCTGCGGTGTTGCACTC
TGGCTGCTGCTTCTGGGCACCGCCGTGTGTATCCACCGCCGGCGCCGAGCTAGGGTGACCTGGGCCCAGGTCTG
TACAGATATACAGTGAGGATGCCATCCTAAACACAGGATGGATCACAGTGACTCCAGTGGTTGGCAGACACT
TGGCGTTCCACCTCTGGCTCTCGGGACCTGAGCAGCAGCAGCAGCCTCAGCAGTCCGGCTGGGGGCGGATGCCCGG
GACCCACTAGACTGTCTGCTCCTTGTCTCTCTGGGACTCCCGAAGCCCCGGCGTGCCCCCTGCTTCCAGACACC
AGCACTTTTTATGGCTCCCTCATCGCTGAGCTGCCCTCCAGTACCCACGCCAGGCCAAGTCCCCAGGTCCCAGCT
GTCAGGCGCCTCCACCCCAGCTGGCCAGCTCTCCAGCCCCCTGTTCCAGCTCAGACAGCCTCTGCAGCCGCAGG
GGACTCTCTTCTCCCCGCTTGTCTCTGGCCCCCTGCAGAGGCTTGGAAAGGCCAAAAAGAAGCAGGAGCTGCAGCAT
GCCAACAGTTCCCCACTGCTCCGGGGCAGCCACTCCTTGGAGCTCCGGGCTGTGAGTTAGGAAATAGAGGTTCC
AAGAACCCTTTCCCAAAGCCAGGAGCTGTGCCCAAGCTCTGGTTGCTTGGCGGGCCCTGGGACCGAAACTCCTC
AGCTCCTCAAATGAGCTGGTTACTCGTCACTCTCCCTCCAGCACCCCTCTTTCCTCATGAAACTCCCCAACTCAG
AGTCAACAGACCCAGCCTCCGGTGGCACCAAGGCTCCCTCCTCCATCCTGCTGCCAGCAGCCCCCATCCCCATC
CTTAGCCCCCTGCAGTCCCCCTAGCCCCCAGGCCTCTTCCCTCTCTGGCCCCAGCCCAAGCTTCCAGTCGCTGTCC
AGCTCCTCACTGTCACTCCTGGGGGAGGATCAAGACAGCGTGTGACCCCTGAGGAGGTAGCCCTGTGCTTGGAA
CTCAGTGAGGGTGAGGAGACTCCAGGAACAGCGTCTCTCCATGCCAAGGGCTCCTTCAACCCCCACCACTAT
GGGTACATCAGCGTCCCAACAGCCTCAGAGTTACGGACATGGGCAGGACTGGAGGAGGGGTGGGGCCCAAGGGG
GGAGTCTTGTGTGCCACCTCGGCCCTGCCTCACCCCCACCCCCAGCGAGGGCTCCTTAGCCAATGGTTGGGGC
TCAGCCTCTGAGGACAATGCCGCCAGCGCCAGAGCCAGCCTTGTGAGTCTCTCCGATGGCTCCTTCCCTCGCTGAT
GCTCACTTTGCCCCGGGCCCTGGCAGTGGCTGTGGATAGCTTTGGTTTCGGTCTAGAGCCCAGGGAGGCAGACTGC
GTCTTCATAGATGCCTCATCACCTCCCTCCCCACGGGATGAGATCTTCTGACCCCCAACCTCTCCCTGCCCTG
TGGGAGTGGAGGCCAGACTGGTTGGAAGACATGGAGGTGAGCCACACCCAGCGGCTGGGAAGGGGGATGCCTCCC
TGGCCCCCTGACTCTCAGATCTCTTCCAGAGAAGTCACTCCACTGTGCTATGCCCAAGGCTGGTGCTTCTCCT
GTAGATTACTCTGAACCGTGTCCCTGAGACTTCCAGACGGGAATCAGAACCCTTCTCCTGTCCACCCACAAG
ACCTGGGCTGTGGTGTGTGGGTCTTGGCCTGTGTTTCTCTGCAGCTGGGGTCCACCTTCCCAAGCCTCCAGAGAG
TTCTCCCTCCACGATTGTGAAAACAAATGAAAACAAATTAGAGCAAAGCTGACCTGGAGCCCTCAGGGAGCAAA
ACATCATCTCCACCTGACTCCTAGCCACTGCTTTCTCCTCTGTGCCATCCACTCCACCAACAGGTTGTTTGGC
CTGAGGAGCAGCCCTGCCTGCTGCTCTTCCCCCACCATTGGATCACAGGAAGTGGAGGAGCCAGAGGTGCCTTT
GTGGAGGACAGCAGTGGCTGCTGGGAGAGGGCTGTGGAGGAAGGAGCTTCTCGGAGCCCCCTCTCAGCCTTACCT
GGGCCCCCTCCTCTAGAGAAGAGCTCAACTCTCTCCCAACCTCACCATGGAAAGAAAAATAATTATGAATGCCACTG
AGGCACTGAGGCCCTACCTCATGCCAAACAAAGGGTTCAAGGCTGGGTCTAGCGAGGATGCTGAAGGAAGGGAGG
TATGAGACCGTAGGTCAAAGCACCATCCTCGTACTGTTGTCACTATGAGCTTAAGAAATTTGATACCATAAAAT
GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 77

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404

<subunit 1 of 1, 985 aa, 1 stop

<MW: 105336, pI: 6.55, NX(S/T): 7

MGGMAQDSPPQILVHPQDQLFQGPGRMSCQASGQPPPTIRWLLNGQPLSMVPPDPHLLP
DGTLLLLQPPARGHAHDGQALSTD LGVYTCEASNRLGTAVSRGARLSVAVLREDFQIQPRDM
VAVVGEQFTLECGPPWGHPEPTVSWWKDGKPLALQGRHTVSGGSLLMARAEKSDEGTYMCV
ATNSAGHRESRAARVSIQEPQDYTEPVELLAVRIQLENVTLNPDPAEGPKPRPAVWLSWKV
SGPAAPAQSYTALFRTQTAPGGQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG
PDSNVLLLLRLPEKVPSAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNLSLPP
ANWTVVGEQTQLEIATHMPGSYCVQVA AVTGAGAGEPSRPVCLLLEQAMERATQEPSEHGWP
TLEQLRATLKRPEVIATCGVALWLLLLGTAVCIHRRRRARVHLGPGLYRYTSEDAILKHRMD
HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRRSLLSWDSRSPGVPLLPDTST
FYGSLIAELPSSTPARPSPQVPAVRRLPPQLAQLSSPCSSSDSLCSRRGLSSPRLSLAPAEA
WKAKKKQELQHANS SPLLRGSHSLELRACELGNRGSKNLSQSPGAVPQALVAWRALGPKLLS
SSNELVTRHLPPAPLFPHETPPTQSQQTQPPVAPQAPSSILLPAAPIPILSPCSPSPQASS
LSGPSPASSRLSSSSSLSSLGEDQDSVLTPEEVALCLELSEGEETPRNSVSPMPRAPSPPTTY
GYISVPTASEFTDMGRTGGGVGPKGGVLLCPPRPCLTPTPSEGLANGWGSASEDNAASARA
SLVSSSDGSFLADAHFARALAVAVDSFGFGLEPREADCVFIDASSPPSPRDEIFLTPNLSLP
LWEWRPDWLEDMEVSHTQRLGRGMPPWPPDSQISSQRSQLHCRMPKAGASPDYS

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristoylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

FIGURE 78

CTCCACGGTGTCCAGCGCCCAGAATGCGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT
CCCAGGTTATGAAGCCCTGGAGGGCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT
CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT
GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTTGTGACCCTGTGGA
ACCTCACCCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG
TCTTTACTGATCTCTCTGTTCGTCTTTCCAGGACCCTGCTGTCTCTCCCTCCCCCTTCTCCAC
CTTCCAGCCTCTGGCTACAACACGCTGCAGCCCAAGGCAAAGCTCAGCAAACCCAGCCCC
CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
GCTGAGGCCCCCTCCATTGCCAGGGACTTCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCCATGCAGC
TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCTGGTGTCTGCTGAGCCTTCTGTGAGC
CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA
CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC
CCTTCCCAGGCCCCCTGAGGGGGACGTGATCTCGATGCCTCCCTCCACACATCTGAGGAGGA
GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT
GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCACCTCAGCCTCAGAG
TCCAGCTGCCCCGACTCCAGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTTGCATGTTCCA
GCCTGACCTAGAAGCGTTTGTGAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG
GAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA
GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTGGGC
CTCATGCCCAGTGTGCGACCCTGCCTTCTCCCTCCACTCCAGACCCACCTTGTCTTCCCTCCC
TGGCGTCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT
GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT
GTGAAAAACGTGATTCTTGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG
GACTCTGAATTCTAACAATGCCAGTGAAGTGTGCACTTGAGTTTGAGGGCCAGTGGGCCTG
ATGAACGCTCACACCCCTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC
CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG
TCCAGGCCTTGGTCAGGTGAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG
TTGCCTTTNCCATTGCCCCCTCCCTGGNCCATGCCTTCTTGCCTTTGGAAAAAATGATGAAGA
AAACCTTGGCTCCTTCTTGTCTGGAAAGGGTTACTTGCCTATGGGTCTGGTGGCTAGAGA
GAAAAGTAGAAAACAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG
ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGCGGGGTGGTGGTAAAGTA
GCACAACTACTATTTTTTTTCTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT
GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACCTCCGCCTCCTGGGTTCAGTGATT
CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT
TTTGTACTTTTAGTAGAGATGGGGTTTTACCATGTTGGCCAGGCTGGTCTTGAACCTCCTGAC
CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG
TCTGGCCCTATTTCTTTTAAAAAGTGAAATTAAGAGTTGTTTCAATATGCAAACTTGGAAG
ATGGAGGAGAAAAAGAAAAGGAAGAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT
TATTTCTGTTTTGTGTAATCTTCTTCCACTTTTTCTTCTTACATAATTTGCCGGTGTCTT
TTTACAGAGCAATTATCTTGTATATACAACCTTGTATCCTGCCTTTTCCACCTTATCGTTCC
ATCACTTTATTCCAGCACTTCTCTGTGTTTTACAGACCTTTTTTATAAATAAAATGTTTCATCA
GCTGCATAAAAAAAAAAAAAA

FIGURE 79

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAE EEGQETMKGRVSI RDSRQELSLIVTLWNLTLDAGEYWCGVEKRGPDESLLISLFV
FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQKGTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSI PMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAE EKEAPSQAPEGD
VISMPPLHTSEEEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

FIGURE 80

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGA
GCCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTC
GCTCTGTTGCCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAA
GCCCTGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTCATCCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAG
CAGGGCTCTCAGAAGGCGGTGGTGGCCAGCTGGGATCATGTTGTTGGCCCTGGTCTGTCTGC
TCAGCTGCCTGCTACCCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTG
CTACATGACTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGC
TTATTTTACAAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTTCCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCCGAACGTCCCCAAC
GTGTGCCGGATGTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGC
CATGAAGATAACCCAAGAGCCTCAGGGTCTGGGTACTGGGAGGCCTGGAGGCATCACTGCC
AGGGAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTCTAGGATGGACGGAACCATGCA
CAGCAGGCTGGGAAATGTGGTTTGGTTCCTGACCTAGGCTTGGGAAGACAAGCCAGCGAATA
AAGGATGGTTGAACGTGAAA

FIGURE 81

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187

<subunit 1 of 1, 146 aa, 1 stop

<MW: 16430, pI: 5.05, NX(S/T): 1

MLLALVCLLSCLLPSSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD

YEADGSTNNGIFQINSRRWCSNLTNPVNPVCRMYCSDLLNPNLKDTVICAMKITQEPQGLGY

WEAWRHHCQGKDLTEWVDGCDF

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homologous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 82

AGCCGCTGCCCCGGGCCGGGCGCCGCGGCGGCACCATGAGTCCCCGCTCGTGCCTGCGTTC
GCTGCGCCTCCTCGTCTTCGCCGTCTTCTCAGCCGCCGCGAGCAACTGGCTGTACCTGGCCA
AGCTGTTCGTTCGGTGGGGAGCATCTCAGAGGAGGAGACGTGCGAGAACTCAAGGGCCTGATC
CAGAGGCAGGTGCAGATGTGCAAGCGGAACCTGGAAGTCATGGACTCGGTGCGCCGCGGTGC
CCAGCTGGCCATTGAGGAGTGCCAGTACCAGTTCCGGAACCGGCGCTGGAAGTGTCCACAC
TCGACTCCTTGCCCGTCTTCGGCAAGGTGGTGACGCAAGGGACTCGGGAGGCGGCCTTCGTG
TACGCCATCTCTTCGGCAGGTGTGGCCTTTGCAGTGACGCGGGCGTGACGAGTGGGGAGCT
GGAGAAGTGCGGCTGTGACAGGACAGTGACATGGGGTCAGCCACAGGGCTTCCAGTGGTCAG
GATGCTCTGACAAACATCGCCTACGGTGTGGCCTTCTCACAGTCGTTTGTGGATGTGCGGGAG
AGAAGCAAGGGGGCCTCGTCCAGCAGAGCCCTCATGAACCTCCACAACAATGAGGCCGGCAG
GAAGGCCATCCTGACACACATGCGGGTGAATGCAAGTGCCACGGGGTGTGAGGCTCCTGTG
AGGTAAAGACGTGCTGGCGAGCCGTGCCGCCCTTCCGCCAGGTGGGTACGCACTGAAGGAG
AAGTTTGATGGTGCCACTGAGGTGGAGCCACGCCGCTGGGCTCCTCCAGGGCACTGGTACC
ACGCAACGCACAGTTCAAGCCGCACACAGATGAGGACCTGGTGTACTTGGAGCCTAGCCCCG
ACTTCTGTGAGCAGGACATGCGCAGCGGCGTGCTGGGCACGAGGGGGCCGCACATGCAACAAG
ACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGCTGTGGCCGCGGCTTCCACACGGCGCA
GGTGGAGCTGGCTGAACGCTGCAGCTGCAAATTCCTACTGGTGCTGCTTCGTCAAGTGCCGGC
AGTGCCAGCGGCTCGTGGAGTTGCACACGTGCCGATGACCGCCTGCCTAGCCCTGCGCCGGC
AACCACCTAGTGGCCCAGGGAAGGCCGATAATTTAAACAGTCTCCACCACCTACCCCAAGA
GATACTGGTTGTATTTTTTGTCTGGTTTGGTTTTTGGGTCCTCATGTTATTTATTGCCGAA
ACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGCCTCCCCAAAGCCTGGGCCTTTGTGGCT
GCCACTGACCAAAGGGACCTTGCTCGTGCCGCTGGCTGCCCCGCATGTGGCTGCCACTGACCA
CTCAGTTGTTATCTGTGTCCGTTTTTCTACTTGCAGACCTAAGGTGGAGTAACAAGGAGTAT
TACCACCACATGGCTACTGACCGTGTGTCATCGGGGAAGAGGGGGCCTTATGGCAGGGAAAATA
GGTACCGACTTGATGGAAGTCACACCTCTGGAAAAAAGAACTCTTAACTCTCCAGCACACA
TACACATGGACTCCTGGCAGCTTGAGCCTAGAAGCCATGTCTCTCAAATGCCCTGAGAAAGG
GAACAAGCAGATACCAGGTCAAGGGCACCAGGTTTATTTAGCCCTTACATGGACAGCTAGA
GGTTTCGATATCTGTGGGTCCTTCCAGGCAAGAAGAGGGAGATGAGAGCAAGAGACGACTGAA
GTCCCACCCTAGAACCCAGCCTGCCCCAGCCTGCCCCTGGAAGAGGAACTTAACCACTCC
CCAGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCCGGCTGTGCCTT
TGCAGTCATGCCCGAGTCACCTTTCACAGCGCTGTTTCTCCATGAAACTGAAAAACACACAC
ACCTGCGAGA
GAGAGGGAGGAAAGGGCTGTGCCCTTTCAGTCATGCCCGAGTCACCTTTCACAGCACTGTTCCCTC

FIGURE 83

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328

<subunit 1 of 1, 351 aa, 1 stop

<MW: 39052, pI: 8.97, NX(S/T): 2

MSPRSCLRSLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVQMCKRNLE
VMDSVRRGAQLAIEECQYQFRNRRWNCSTLDSL PVFGKVVTQGTREAAFVYAISSAGVAFV
TRACSSGELEKCGCDRTVHGVSPQGFQWSGCSDNIAYGVAFSQSFVDVRERSKGASSSRALM
NLHNNEAGRKAILTHMRVECKCHGVSGSEVKTCWRAVPPFRQVGHALKEKFDGATEVEPRR
VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAIDGCELLC
CGRGFHTAQVELAERCSCKFHWCCFVKCRQCQRLVELHTCR

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 88-91 and 297-300

Wnt-1 family signature.

amino acids 206-215

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCAT
CGCCATGGACACCACCAGGTACAGCAAGTGGGGCGGCAGCTCCGAGGAGGTCCCCGGAGGGC
CCTGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCTTGGTC
ACCACAGTCCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGC
GGCGCTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCCTCGAAGCAGACGGCGGCGCTGG
GTGCCCTGAAGGAGGAGGTCTGGAGACTGCCACAGCTGCTGCTCGGGGACGCAGGCGCAGCTG
CAGACCACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCT
GCGGGAAGTGCCTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGGCCGTGAGGACG
TCCGCACTGAGCTGTTCCGGGCGCTGGAGGGCCGTGAGGCTCCAGAACAACCTCTGCGAGCCG
TGCCCCACGTCGTGGCTGTCTTTCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGAC
GTGGGCGGCGGCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGCC
TGGATGAGCAGGGCTTCCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCCTGAGG
GCTGTGCGCCATCTGGGCAAGGTTCAAGGCTACCAGTGGGTGGACGGAGTCTCTCTCAGCTT
CAGCCACTGGAACCAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAACTGTGTCATGATGC
TGCACACGGGGCTGTGGAACGACGCACCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAG
AAAAGGCACAACTGCTTGACCCCGCCAGTGCCCTGGAGCCGCGCCCATTGCAGCATGTCGTA
TCCTGGGGGCTGCTCACCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTCTTCCT
CATCCACCGCTGCTGAGTCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAGTGCC
TGGGCTCTGGGACCTCCATGCCGACCTCATCCTAACTCCACTCACGCAGACCCAACCTAACC
TCCACTAGCTCCAAAATCCCTGCTCCTGCGTCCCCGTGATATGCCTCCACTTCTCTCCCTAA
CCAAGGTTAGGTGACTGAGGACTGGAGCTGTTTGGTTTTCTCGCATTTTCCACCAAACCTGGA
AGCTGTTTTTGCAGCCTGAGGAAGCATCAATAAATATTTGAGAAATGAAAAA

FIGURE 85

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352

<subunit 1 of 1, 293 aa, 1 stop

<MW: 32562, pI: 6.53, NX(S/T): 2

MDTTRYSKWGGSSSEVPGGPWGRWVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAA
LLDGHDLLRTNASKQTAALGALKEEVGDCHSCCSGTQAQLQTTRAELGEAQAKLMEQESALR
ELRERVVTQGLAEAGRGREDVRTELFRALAVRLQNNSCPCPTSWLSFEGSCYFFSVPKTTW
AAAQDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSFS
HWNQGEPNDAWGRENVCVMMMLHTGLWNDAPCDSEKDGWICEKRHNC

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 86

GCCAGGGGAAGAGGGTGATCCGACCCGGGGAAGGTCGCTGGGCAGGGCGAGTTGGGAAAGCG
GCAGCCCCCGCCGCCCCCGCAGCCCCCTTCTCCTCCTTTCTCCCACGTCCTATCTGCCTCTCG
CTGGAGGCCAGGCCGTGCAGCATCGAAGACAGGAGGAACCTGGAGCCTCATTGGCCGGCCCCGG
GGCGCCGGCCTCGGGCTTAAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACCGCTGCCGGC
CGCGCTCCCGCTGCTCCTGCCGGGTGATGGAAAACCCAGCCCCGGCCGCGCCCTGGGCAAG
GCCCTCTGCGCTCTCCTCCTGGCCACTCTCGGCGCCGCGGCCAGCCTCTTGGGGGAGAGTC
CATCTGTTCCGCCAGAGCCCCGGCCAAATACAGCATCACCTTCACGGGCAAGTGAGCCAGA
CGGCCTTCCCCAAGCAGTACCCCCTGTTCCGCCCCCTGCGCAGTGGTCTTCGCTGCTGGGG
GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACCAGTACGTACGTAAACGGGCTGCG
CGACTTTGCGGAGCGCGGCGAGGCCTGGGCGCTGATGAAGGAGATCGAGGCGGCGGGGGAGG
CGCTGCAGAGCGTGCACGAGGTGTTTTTCGGCGCCCGCCGTCCCCAGCGGCACCGGGCAGACG
TCGGCGGAGCTGGAGGTGCAGCGCAGGCACTCGCTGGTCTCGTTTGTGGTGCGCATCGTGCC
CAGCCCCGACTGGTTTCGTGGGCGTGGACAGCCTGGACCTGTGCGACGGGGACCGTTGGCGGG
AACAGGCGGCGCTGGACCTGTACCCCTACGACGCGGGACGGACAGCGGCTTCACCTTCTCC
TCCCCCAACTTCGCCACCATCCCGCAGGACACGGTGACCGAGATAACGTCCTCCTCTCCCAG
CCACCCGGCCAACTCCTTCTACTACCCGCGGCTGAAGGCCCTGCCTCCCATCGCCAGGGTGA
CACTGCTGCGGCTGCGACAGAGCCCCAGGGCCTTCATCCCTCCCGCCCCAGTCTTGCCACG
AGGGACAATGAGATTGTAGACAGCGCCTCAGTTCCAGAAACGCCGCTGGACTGCGAGGTCTC
CCTGTGGTTCGTCTTGGGACTGTGCGGAGGCCACTGTGGGAGGCTCGGGACCAAGAGCAGGA
CTCGCTACGTCCGGGTCCAGCCCGCCAACAACGGGAGCCCCCTGCCCCGAGCTCGAAGAAGAG
GCTGAGTGCGTCCCTGATAACTGCGTCTAAGACCAGAGCCCCCGCAGCCCCCTGGGGCCCCCG
GAGCCATGGGGTGTGCGGGGCTCCTGTGCAGGCTCATGCTGCAGGCGGCCGAGGGCACAGGG
GGTTTCGCGCTGCTCCTGACCGCGGTGAGGCCGCGCCGACCATCTCTGCACTGAAGGGCCCT
CTGGTGGCCGGCACGGGCATTGGGAAACAGCCTCCTCCTTTCCCAACCTTGCTTCTTAGGGG
CCCCCGTGTCCCGTCTGCTCTCAGCCTCCTCCTCCTGCAGGATAAAGTCATCCCCAAGGCTC
CAGCTACTCTAAATTATGTCTCCTTATAAGTTATTGCTGCTCCAGGAGATTGTCCTTCATCG
TCCAGGGGCCTGGCTCCACGTGGTTGCAGATACCTCAGACCTGGTGTCTAGGCTGTGCTG
AGCCCACTCTCCCGAGGGCGCATCCAAGCGGGGGCCACTTGAGAAGTGAATAAATGGGGCGG
TTTCGGAAGCGTCAGTGTTCATGTTATGGATCTCTCTGCGTTTGAATAAAGACTATCTCT
GTTGCTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 87

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53971

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35844, pI: 5.45, NX(S/T): 2

MENPSPAAALGKALCALLLATLGAAGQPLGGESICSARAPAKYSITFTGKWSQTAFPKQYPL
FRPPAQWSSLLGAAHSSDYSMWRKNQYVSNGLRDFAEERGEAWALMKEIEAAGEALQSVHEVF
SAPAVPSGTGQTSAELEVQRRHSLVSFVVRIVPSPDWFGVDSLCLCDGDRWREQAALDLYP
YDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHPANSFYYPRLKALPPIARVTLLRLRQSP
RAFIPPAPVLPSRDNEIVDSASVPETPLDCEVSLWSSWGLCGGHCGR LGTKSRTRYVRVQPA
NNGSPCPELEEEEAECVPDNCV

Important features:

Signal peptide:

amino acids 1-26

FIGURE 88

GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTAGTGTGTTGGTGTCCCTGTCTTGCGTGA
TATTGACAAACTGAAGCTTTCTGCACCACTGGACTTAAGGAAGAGTGTA CTCTAGGCGGA
CAGCTTTAGTGGCCGGCCGGCCGCTCTCATCCCCGTAAGGAGCAGAGTCCTTTGTACTGAC
CAAGATGAGCAACATCTACATCCAGGAGCCTCCACGAATGGGAAGGTTTTATTGAAAATA
CAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAGAAGCTCCTAAAGCTTGCAGAAATTTT
ATCCAACCTTTGTTTGGGAAGCTTATTATGACAATACCATTTTTCATAGAGTTGTGCCTGGTTT
CATAGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTCTATCTATGGAGCGC
CATTCAAAGATGAATTTCAATCACGGTTGCGTTTTAATCGGAGAGGACTGGTTGCCATGGCA
AATGCTGGTTCTCATGATAATGGCAGCCAGTTTTTCTTCACACTGGGTGAGCAGATGAACT
TAACAATAAGCATAACCATCTTTGGAAAGGTTACAGGGGATACAGTATATAACATGTTGCGAC
TGTCAGAAGTAGACATTGATGATGACGAAAGACCACATAATCCACACAAAATAAAAAGCTGT
GAGGTTTTGTTTAATCCTTTTGATGACATCATTTCCAAGGGAAATTAAGGCTGAAAAAGA
GAAACCAGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTTAGTTTACTTT
CATTTGGAGAGGAAGCTGAGGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAG
GGCAAAAGCAAAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGT
TG TAGAAAGTGAAAAAGGTGATGCACCAGATTTAGTTGATGATGGAGAAGATGAAAGTGCAG
AGCATGATGAATATATTGATGGTGATGAAAAGAACCTGATGAGAGAAAGAATTGCCAAAAAA
TTAAAAAAGGACACAAGTGCGAATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATC
AGTCAGCCGCAGTGAAGAGCTCAGAAAAGAAGCAAGACAATTAAAACGGGAACTCTTAGCAG
CAAAACAAAAAAAAGTAGAAAATGCAGCAAAACAAGCAGAAAAAAGAAGTGAAGAGGAAGAA
GCCCCCTCAGATGGTGCTGTTGCCGAATACAGAAGAGAAAAGCAAAAGTATGAAGCTTTGAG
GAAGCAACAGTCAAAGAAGGGAACCTCCCGGGAAGATCAGACCTTGCACTGCTGAACCAGT
TTAAATCTAACTCACTCAAGCAATTGCTGAAACACCTGAAAATGACATTCCTGAAACAGAA
GTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTTGAGGATAAAAGCAGAAAAGT
GAAAGATGCAAGCATGCAAGACTCAGATACATTTGAAATCTATGATCCTCGGAATCCAGTGA
ATAAAAGAAGGAGGGAAGAAAGCAAAAAGCTGATGAGAGAGAAAAAAGAAAGAAGATAAAT
GAGAATAATGATAACCAGAACTTGCTGGAAATGTGCCTACAATGGCCTTGTAACAGCCATTG
TTCCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTTGAACCTGTTGTCTGGTTTTG
AAAAACAATTATCTTGTTTTGCAAATTGTGGAATGATGTAAGCAAATGCTTTTGGTTACTGG
TACATGTGTTTTTTCTAGCTGACCTTTTATATTGCTAAATCTGAAATAAAATAACTTTCCT
TCCACAAAAA

FIGURE 89

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919

><subunit 1 of 1, 472 aa, 1 stop

><MW: 53847, pI: 5.75, NX(S/T): 2

MSNIYIQEPPTNGKVLLKTTAGDIDIELWSKEAPKACRNFIQLCLEAYYDNTIFHRVVPGFI
VQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRADELN
NKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPHKIKSCEVLFNPFDDIIPREIKRLKKEK
PEEEVKKLKPKGTKNFSLLSFGEAAEEEEEEVNVRVSQSMKGKSKSSHDLKDDPHLSSVPV
ESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKKLKKDTSANVKSAGEGEVEKKSV
SRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEEAPPDGAVAfYRREKQKYEALRK
QQSKKGTSREDQTLALLNQFKSKLTQAIAETPENDIPETEVEDDEGWMSHVLQFEDKSRKVK
DASMQSDTFEIYDPRNPVNKRREESKKLMREKKERR

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.

amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase

amino acids 96-140, 49-89 and 22-51

FIGURE 90

CGCCGCCGTTGGGGCTGGAAGTTCCCGCCAGGTCCGTGCCGGGCGAGAGAGATGCTGCCCGG
CCCGCCTCGGCTTTGAGGCGAGAGAAGTGTCCCAGACCCATTTTCGCCCTTGCTGACGGCGTCCG
AGCCCTGGCCAGACATGTCCACAGGGTTCTCCTTCGGGTCCGGGACTCTGGGCTCCACCACC
GTGGCCCGCGGCGGGACCAGCACAGGCGGCGTTTTCTCCTTCGGAACGGGAACGTCTAGCAA
CCCTTCTGTGGGGCTCAATTTTGAAATCTTGGAAGTACTTCAACTCCAGCAACTACATCTG
CTCCTTCAAGTGGTTTTTGAACCGGGCTCTTTGGATCTAAACCTGCCACTGGGTTCACTCTA
GGAGGAACAAATACAGGTGCCTTGCACACCAAGAGGCCTCAAGTGGTACCAAATATGGAAC
CCTGCAAGGAAAACAGATGCATGTGGGGAAGACACCCATCCAAGTCTTTTTTAGGAGTCCCCT
TCTCCAGACCTCCTCTAGGTATCCTCAGGTTTGCACCTCCAGAACCCCGGAGCCCTGGA
GGAATCAGAGATGCTACCACCTACCCGCCTGGATGGAGTCTCGCTCTGTGCGCAGGCTGGAG
TGCAGTGGCAGCATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGAGTCTCCTGC
CTCAGCCTCTGAGTGTCTGGGGCTACAGGTGCCTGCAGGAGTCTTGGGGCCAGCTGGCCTCG
ATGTACGTACGACGCGGGAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCT
GAACGTGTACGCGCCGGCGCGCGCGCCCGGGGATCCCAGCTGCCAGTGATGGTCTGGTTCC
CGGGAGGCGCCTTCATCGTGGGCGCTGCTTCTTCGTACGAGGGCTCTGACTTGGCCGCCCGC
GAGAAAGTGGTGTCTGGTGTCTTCTGCAGCACAGGCTCGGCATCTTCGGCTTCCTGAGCACGGA
CGACAGCCACGCGCGCGGGAACCTGGGGGCTGCTGGACCAGATGGCGGCTCTGCGCTGGGTGC
AGGAGAACATCGCAGCCTTCGGGGGAGACCCAGGAAATGTGACCTGTTTCGGCCAGTCCGGC
GGGGCCATGAGCATCTCAGGACTGATGATGTACCCCTAGCCTCGGGTCTCTTCCATCGGGC
CATTTCCAGAGTGGCACCGCGTTATTACAGACTTTTCATCACTAGTAACCCACTGAAAGTGG
CCAAGAAGGTTGCCACCTGGCTGGATGCAACCACAACAGCACACAGATCCTGGTAAACTGC
CTGAGGGCACTATCAGGGACCAAGGTGATGCGTGTGTCCAACAAGATGAGATTCTTCCA
GAACTTCCAGAGAGACCCGGAAGAGATTATCTGGTCCATGAGCCCTGTGGTGGATGGTGTGG
TGATCCCAGATGACCTTTGGTGCTCCTGACCCAGGGGAAGGTTTCATCTGTGCCCTACCTT
CTAGGTGTCAACAACCTGGAATTCAATTGGCTCTTGCCCTTATAATATACCAAGGAGCAGGT
ACCACTTGTGGTGGAGGAGTACCTGGACAATGTCAATGAGCATGACTGGAAGATGCTACGAA
ACCGTATGATGGACATAGTTCAAGATGCCACTTTCGTGTATGCCACACTGCAGACTGCTCAC
TACCACCGAGAAACCCCAATGATGGGAATCTGCCCTGCTGGCCACGCTACAACAAGGATGAA
AAGTACCTGCAGCTGGATTTTACCACAAGAGTGGGCATGAAGCTCAAGGAGAAGAAGATGGC
TTTTTGGATGAGTCTGTACCAGTCTCAAAGACCTGAGAAGCAGAGGCAATTCTAAGGGTGGC
TATGCAGGAAGGAGCCAAAGAGGGGTTTGCCCCCACCATCCAGGCCCTGGGGAGACTAGCCA
TGGACATACCTGGGGACAAGAGTTCTACCCACCCAGTTTAGAACTGCAGGAGCTCCCTGCT
GCCTCCAGGCCAAAGCTAGAGCTTTTGCCTGTTGTGTGGGACCTGCACTGCCCTTTCCAGCC
TGACATCCCATGATGCCCCCTCTACTTCACTGTTGACATCCAGTTAGGCCAGGCCCTGTCAAC
ACCACACTGTGCTCAGCTCTCCAGCCTCAGGACAACCTCTTTTTTTTCCCTTCTTCAAATCCT
CCCACCCTTCAATGTCTCCTTGTGACTCCTTCTTATGGGAGGTCGACCCAGACTGCCACTGC
CCCTGTCACTGCACCCAGCTTGGCATTACCATCCATCCTGCTCAACCTTGTTCCTGTCTGT
TCACATTGGCCTGGAGGCCTAGGGCAGGTTGTGACATGGAGCAAACCTTTTGGTAGTTTGGGA
TCTTCTCTCCACCCACACTTATCTCCCCAGGGCCACTCCAAAGTCTATACACAGGGGTGG
TCTCTTCAATAAAGAAGTGTTGATTAGAAAAA

Table 1. The 1977-1978 influenza season	
Age group	Number of cases
0-4	100
5-9	100
10-14	100
15-19	100
20-24	100
25-29	100
30-34	100
35-39	100
40-44	100
45-49	100
50-54	100
55-59	100
60-64	100
65-69	100
70-74	100
75-79	100
80-84	100
85-89	100
90-94	100
95-99	100
100	100

amino acids 318-321, 380-383 and 465-468

FIGURE 93

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002

><subunit 1 of 1, 544 aa, 1 stop

><MW: 60268, pI: 9.53, NX(S/T): 3

MLLPLLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFSPYPRQDWTGSTPAYGYWFK
AVTETTTKGAPVATNHQSREVESTRGRFQLTGDPKGNCSLVIRDAQMQDESQYFFRVERGS
YVTYNFMNDGFFLKVTVLSFTPRPQDHNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLCAADSQPPATLSWVLQNRVLSSSHWPWGPRPL
GLELPGVKAGDSGRYTCRAENRLGSQQRALDLSVQYPPENLRVMVSQANRTVLENLGNGTSL
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSQSPSDPGVLELPRVQVEHEGEFTCHARHP
LGSQHVSLSLSVHYKKGLISTAFSNGAFLGIGITALLFLCLALIIMKILPKRRTQTETPRPR
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPPGAPSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVRPRPEARMPKGTQADYAEVKFQ

Important features:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 100-103, 297-300 and 306-309

**Immunoglobulins and major histocompatibility complex proteins
signature.**

amino acids 365-371

FIGURE 95

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906

><subunit 1 of 1, 772 aa, 1 stop

><MW: 87002, pI: 4.64, NX(S/T): 8

MNCYLLLRFMLGIPLLWPCLGATENSQTKKVKQPVRSHLRVKGWVWNQFFVPEEMNTTSHH
IGQLRSDLDNGNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVSNDINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL
LYSLAQGPYPYFSVEPTTGIVIRISSKMDRELQDEYWVIIQAKDMIGQPGALSGTTSVLIKLSL
VNDNKPIFKESLYRLTVSESAPTGTSGTITMAYDNDIGENAEMDYSIEEDDSQTFDIIITNHE
TQEGIVILKKKVD FEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFL
PYYVFEVFEETPQGSFVGVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR
DESIEEHFHYFNLSVEDTNNSSFTIIDNQDNTAVILTNRTGFNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVCDGSGSTQTCQYQELVLSMGFKTEVIIAILICIMIIFGFIFLTLGLKQ
RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTSAEIRSLY
RQSLQVGPDSAIFRKFILEKLEEANTDPCAPPFDSLQTYAFEGTGSLAGSLSSLESASVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSNN

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518,
516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

FIGURE 96

ATTTCAAGGCCAGCCATATTTTTNTGTTGAACCAACAACAGGAGTCATAAGAATATTTNTA
AAATGGATAGAGAACTGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGT
CAGCCAGGAGCGTTGTNTGGAACAACAAGTGTATTAATTAACTTTCAGATGTTAATGACAA
TAAGCCTATATTTAAAGAAAGTTTATACCGCTTGACTGTNTNTGAATCTGCACCCACTGGGA
NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC
AGCATTGAAGAGGATGATTCGCAAACATTTGACATTATT

FIGURE 97

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCGGGCGCGGACCCCAACCCCGAC
CCAGAGCTTCTCCAGCGGCGGCGCAGCGAGCAGGGCTCCCCGCTTAACTTCCTCCGCGGGG
CCCAGCCACCTTCGGGAGTCCGGGTTGCCACCTGCAAACCTCTCCGCCTTCTGCACCTGCCA
CCCCTGAGCCAGCGCGGGCCCCCGAGCGAGTCAATGGCCAACGCGGGGCTGCAGCTGTTGGGC
TTCATTCTCGCCTTCCTGGGATGGATCGGCGCCATCGTCAGCACTGCCCTGCCCCAGTGGAG
GATTTACTCCTATGCCGGCGACAACATCGTGACCGCCAGGCCATGTACGAGGGGCTGTGGA
TGTCCTGCGTGTGCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAAT
CTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTGGCATCCTCCTGGGAGTGAT
AGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGC
AGAAGATGAGGATGGCTGTCAATTGGGGGTGCGATATTTCTTCTTGCAAGTCTGGCTATTTTA
GTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCATGACCCTATGACCCCAGT
CAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCC
TTCTGGGAGGTGCCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCAACACCA
AGGCCCTATCCAAAACCTGCACCTTCCAGCGGGAAAGACTACGTGTGACACAGAGGCAAAAG
GAGAAAATCATGTTGAAACAAACCGAAAATGGACATTGAGATACTATCATTAACATTAGGAC
CTTAGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAACAAACAAACAAACAAAA
ACCCATGTGTTAAAATACTCAGTGCTAAACATGGCTTAATCTTATTTTATCTTCTTTCCTCA
ATATAGGAGGGAAGATTTTTCATTTGTATTACTGCTTCCATTGAGTAATCATACTCAAAT
GGGGGAAGGGGTGCTCCTTAAATATATATAGATATGTATATATACATGTTTTTCTATTAAAA
ATAGACAGTAAAATACTATTCTCATTATGTTGATACTAGCATACTTAAAATATCTCTAAAT
AGGTAAATGTATTTAATTCATATTGATGAAGATGTTTATTGGTATATTTTCTTTTTCGTCC
TTATATACATATGTAACAGTCAAATATCATTTACTCTTCTTCATTAGCTTTGGGTGCCTTTG
CCACAAGACCTAGCCTAATTTACCAAGGATGAATTCCTTCAATTCTTCATGCGTGCCCTTTT
CATATACTTATTTTATTTTACCATAATCTTATAGCACTTGCATCGTTATTAAGCCCTTAT
TTGTTTTGTGTTTCATTGGTCTCTATCTCCTGAATCTAACACATTTCATAGCCTACATTTTA
GTTTCTAAAGCCAAGAAGAATTTATTACAAATCAGAACCTTGGAGGCAAATCTTTCTGCATG
ACCAAAGTGATAAATTCCTGTTGACCTTCCCACACAATCCCTGTACTCTGACCCATAGCACT
CTTGTTTGTCTTTGAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTCCCCAGGTGTTGT
AACACAACCTTTATTTGATTGAATTTTTAAGCTACTTATTCATAGTTTTATATCCCCCTAACT
ACCTTTTTGTTCCTTAAATTTGTATTGTTTTCCCAAGTGTAATTATCATGCGTTTTTA
TATCTTCCTAATAAGGTGTGGTCTGTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATA
ATCTGGTGACAAATATTCTCTCTGTAGCTGTAAGCAAGTCACTTAATCTTTCTACCTCTTTT
TTCTATCTGCCAAATTGAGATAATGATACTTAACCAGTTAGAAGAGGTAGTGTGAATATTAA
TTAGTTTATATTACTCTTATTCTTTGAACATGAACTATGCCTATGTAGTGTCTTTATTTGCT
CAGCTGGCTGAGACACTGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCATGTGATT
CACTGCCTTCCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACACACATACCTTCAT
GTGGTTCAAGTGCCTTCCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACGCACATAC
CTTCATGTGGCTCAGTGCCTTCCTCTCTCTACCAGTCTATTTCCATTCTTTCACTGTGTCT
GACATGTTTGTGCTCTGTTCCATTTTAACAACCTGCTCTTACTTTTCCAGTCTGTACAGAATG
CTATTTCACTTGAGCAAGATGATGTAATGGAAAGGGTGTGGCACTGGTGTCTGGAGACCTG
GATTTGAGTCTTGGTGCTATCAATCACCGTCTGTGTTTGGAGCAAGGCATTTGGCTGCTGTAA
GCTTATTGCTTCATCTGTAAGCGGTGGTTTGTAAATTCCTGATCTTCCCACCTCACAGTGATG
TTGTGGGGATCCAGTGAGATAGAATACATGTAAGTGTGGTTTTGTAAATTTAAAAGTGCTAT
ACTAAGGGAAAGAATTGAGGAATTAACGCATACGTTTTGGTGTGCTTTTCAAATGTTTGA
AAATAAAAAAATGTTAAG

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185

><subunit 1 of 1, 211 aa, 1 stop

><MW: 22744, pI: 8.51, NX(S/T): 1

MANAGLQLLGFI~~L~~AF~~L~~GWIGAI~~V~~STALPQWRIYSYAGDNIVTAQAMYEGLWMSCVSQSTGQI
QCKVFDSL~~L~~NLSSTLQATRALMVVGILLGVIAIFVATVGMKCMKCLEDDDEVQKMRMAVIGGA
IFLLAGLAILVATAWYGNRIVQEFYDPMPV~~N~~ARYEFGQALFTGWAAASLCLLGALLCCSC
PRKTT~~S~~YPTPRPYPKPAPSSSGKD~~Y~~V

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-102, 118-142 and 161-187

N-glycosylation site.

amino acids 72-75

PMP-22 / EMP / MP20 family proteins

amino acids 70-111

ABC-2 type transport system integral membrane protein

amino acids 119-133

Parameter	Value	Unit	Source
α	0.001	cm ² s ⁻¹	Table 1
β	0.001	cm ² s ⁻¹	Table 1
γ	0.001	cm ² s ⁻¹	Table 1
δ	0.001	cm ² s ⁻¹	Table 1
ϵ	0.001	cm ² s ⁻¹	Table 1
ζ	0.001	cm ² s ⁻¹	Table 1
η	0.001	cm ² s ⁻¹	Table 1
θ	0.001	cm ² s ⁻¹	Table 1
ι	0.001	cm ² s ⁻¹	Table 1
κ	0.001	cm ² s ⁻¹	Table 1
λ	0.001	cm ² s ⁻¹	Table 1
μ	0.001	cm ² s ⁻¹	Table 1
ν	0.001	cm ² s ⁻¹	Table 1
ξ	0.001	cm ² s ⁻¹	Table 1
\omicron	0.001	cm ² s ⁻¹	Table 1
π	0.001	cm ² s ⁻¹	Table 1
ρ	0.001	cm ² s ⁻¹	Table 1
σ	0.001	cm ² s ⁻¹	Table 1
τ	0.001	cm ² s ⁻¹	Table 1
υ	0.001	cm ² s ⁻¹	Table 1
ϕ	0.001	cm ² s ⁻¹	Table 1
χ	0.001	cm ² s ⁻¹	Table 1
ψ	0.001	cm ² s ⁻¹	Table 1
ω	0.001	cm ² s ⁻¹	Table 1
φ	0.001	cm ² s ⁻¹	Table 1
ϑ	0.001	cm ² s ⁻¹	Table 1
ϖ	0.001	cm ² s ⁻¹	Table 1
ς	0.001	cm ² s ⁻¹	Table 1
η	0.001	cm ² s ⁻¹	Table 1
θ	0.001	cm ² s ⁻¹	Table 1
ι	0.001	cm ² s ⁻¹	Table 1
κ	0.001	cm ² s ⁻¹	Table 1
λ	0.001	cm ² s ⁻¹	Table 1
μ	0.001	cm ² s ⁻¹	Table 1
ν	0.001	cm ² s ⁻¹	Table 1
ξ	0.001	cm ² s ⁻¹	Table 1
\omicron	0.001	cm ² s ⁻¹	Table 1
π	0.001	cm ² s ⁻¹	Table 1
ρ	0.001	cm ² s ⁻¹	Table 1
σ	0.001	cm ² s ⁻¹	Table 1
τ	0.001	cm ² s ⁻¹	Table 1
υ	0.001	cm ² s ⁻¹	Table 1
ϕ	0.001	cm ² s ⁻¹	Table 1
χ	0.001	cm ² s ⁻¹	Table 1
ψ	0.001	cm ² s ⁻¹	Table 1
ω	0.001	cm ² s ⁻¹	Table 1
φ	0.001	cm ² s ⁻¹	Table 1
ϑ	0.001	cm ² s ⁻¹	Table 1
ϖ	0.001	cm ² s ⁻¹	Table 1
ς	0.001	cm ² s ⁻¹	Table 1
η	0.001	cm ² s ⁻¹	Table 1
θ	0.001	cm ² s ⁻¹	Table 1
ι	0.001	cm ² s ⁻¹	Table 1
κ	0.001	cm ² s ⁻¹	Table 1
λ	0.001	cm ² s ⁻¹	Table 1
μ	0.001	cm ² s ⁻¹	Table 1
ν	0.001	cm ² s ⁻¹	Table 1
ξ	0.001	cm ² s ⁻¹	Table 1
\omicron	0.001	cm ² s ⁻¹	Table 1
π	0.001	cm ² s ⁻¹	Table 1
ρ	0.001	cm ² s ⁻¹	Table 1
σ	0.001	cm ² s ⁻¹	Table 1
τ	0.001	cm ² s ⁻¹	Table 1
υ	0.001	cm ² s ⁻¹	Table 1
ϕ	0.001	cm ² s ⁻¹	Table 1
χ	0.001	cm ² s ⁻¹	Table 1
ψ	0.001	cm ² s ⁻¹	Table 1
ω	0.001	cm ² s ⁻¹	Table 1
φ			

ACCCTTGACCCAACGCGGCCCGGACCGNNTTCATGGCCAAACGCGGGNCTCCAGCTGTTGG
 GCTTCATTCTCCCCTTCTGGGATGGACCGGCGCCCATCNTCAGCACTGCCCTGCCCCAGTG
 GAGGATTTACTCCTATNCCGGCNACAACATCGTGACCGCCAGGCCNTGTACGAGGGGCTGT
 GGATGTCCTGCGTGTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCCTTGCT
 GAATCTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAG
 TGATAGCAATCTTNNTGGCCACCGTTGTNNNTGAAGTGTATGAAGTGCTTGGAAGACGATGA
 GGTGCAGAAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGACGGTCTGGCTA
 TTTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCGA

FIGURE 101

GGGCCC GACCATTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC
CATCNTCAGACTCCCTGCCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGGCTTTGAACGTCNGCGTGTGCGAGANCACCGGGCAGATCCAGTGCAA
AGTCTTTGACTCCTTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGT
GCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTT
CTTGCAGGTCTGGCTATTTNNNGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT
TCTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGC
TGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTTCCTGCGA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Parameter	Value	Unit
Initial concentration of H_2O_2	0.01	M
Initial concentration of Fe^{2+}	0.001	M
Initial concentration of H^+	0.1	M
Temperature	25	$^{\circ}\text{C}$
Reaction time	0-100	min
Reaction rate constant k	0.001	min^{-1}
Reaction order	1	
Half-life $t_{1/2}$	69.3	min
Activation energy E_a	10.5	kJ/mol
Pre-exponential factor A	0.001	min^{-1}
Reaction mechanism	Free radical chain reaction	
Rate-determining step	Decomposition of Fe^{2+}	
Overall reaction	$\text{H}_2\text{O}_2 \rightarrow \text{H}_2\text{O} + \text{O}_2$	
Yield of O_2	1.0	
Yield of H_2O	1.0	
Yield of Fe^{3+}	0.001	M
Yield of H^+	0.1	M
Yield of OH^-	0.001	M
Yield of Fe^{2+}	0.001	M
Yield of H_2O_2	0.01	M
Yield of O_2	1.0	
Yield of H_2O	1.0	
Yield of Fe^{3+}	0.001	M
Yield of H^+	0.1	M
Yield of OH^-	0.001	M
Yield of Fe^{2+}	0.001	M
Yield of H_2O_2	0.01	M
Yield of O_2	1.0	
Yield of H_2O	1.0	
Yield of Fe^{3+}	0.001	M
Yield of H^+	0.1	M
Yield of OH^-	0.001	M
Yield of Fe^{2+}	0.001	M
Yield of H_2O_2	0.01	M
Yield of O_2	1.0	
Yield of H_2O	1.0	
Yield of Fe^{3+}	0.001	M
Yield of H^+	0.1	M
Yield of OH^-	0.001	M
Yield of Fe^{2+}	0.001	M
Yield of H_2O_2	0.01	M
Yield of O_2	1.0	
Yield of H_2O	1.0	
Yield of Fe^{3+}	0.001	M
Yield of H^+	0.1	M
Yield of OH^-	0.001	M
Yield of Fe^{2+}	0.001	M
Yield of H_2O_2	0.01	M
Yield of O_2	1.0	
Yield of H_2O	1.0	
Yield of Fe^{3+}	0.001	M
Yield of H^+	0.1	M
Yield of OH^-	0.001	M
Yield of Fe^{2+}	0.001	M
Yield of H_2O_2	0.01	M
Yield of O_2	1.0	
Yield of H_2O	1.0	
Yield of Fe^{3+}	0.001	M
Yield of H^+	0.1	M
Yield of OH^-	0.001	M
Yield of Fe^{2+}	0.001	M
Yield of H_2O_2	0.01	M
Yield of O_2	1.0	
Yield of H_2O	1.0	
Yield of Fe^{3+}	0.001	M
Yield of H^+	0.1	M
Yield of OH^-	0.001	M
Yield of Fe^{2+}	0.001	M
Yield of H_2O_2	0.01	M
Yield of O_2	1.0	
Yield of H_2O	1.0	
Yield of Fe^{3+}	0.001	M
Yield of H^+	0.1	M
Yield of OH^-	0.001	M
Yield of Fe^{2+}	0.001	M
Yield of H_2O_2	0.01	M
Yield of O_2	1.0	
Yield of H_2O	1.0	
Yield of Fe^{3+}	0.001	M
Yield of H^+	0.1	M
Yield of OH^-	0.001	M
Yield of Fe^{2+}	0.001	M
Yield of H_2O_2	0.01	M
Yield of O_2	1.0	
Yield of H_2O	1.0	
Yield of Fe^{3+}	0.001	M
Yield of H^+	0.1	M
Yield of OH^-	0.001	M
Yield of Fe^{2+}	0.001	M
Yield of H_2O_2	0.01	M
Yield of O_2	1.0	
Yield of H_2O	1.0	
Yield of Fe^{3+}	0.001	M
Yield of $\$		

ATTCTCCCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTGACCCCCCAGGCCATTTACCGAGGGGCTTTGGATGTCNT
GCNTGTGCGAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG
CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCCTCCTGGGAGTGATAGCAAC
CTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCAATTGGGGGCGCGATATTTCTTGTTGCAGGTCTGGCTATTTTAGTNGC
CACAGCATGGTATGGCAATAGANTNNTTCNNGNNNTCTATGACCCTATGACCCCAGTCAATG
CCAGGTACGAATTTGGTCAAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTG
GGAGGTGCCCTACTTTGCTGTTCCCTGTCCC

FIGURE 103

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTTGACCCTTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTTGCCCTTGAAGGTGGTTGNCATCCCCCCTGGGAGTGAATAGCAATCTTTGTG
GCCACCGTTGGCATGAAGTNTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGTNCCACAGCAT
GGTATGGCAATAGNATNNTTCGNGGNTTCTATGACCCTATGACCCCAGTCAATGCCAGGTAC
GAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGC
CCTACTTTGCTGTTCTGTCCCCGAA

FIGURE 105

TCATAGGGGGGCGCGATATTTTTCTTGACGGTNTGGTTATTTTAGTTGCCACAGCATGGTA
TGGCAATAGAAATCGTTCAAGAATTNTATGACCCTATGACCCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA
CTTTGCTGTTTCCTG

TCATAGGGGGGCGCGATATTTTTCTTGACGGTNTGGTTATTTTAGTTGCCACAGCATGGTA
TGGCAATAGAAATCGTTCAAGAATTNTATGACCCTATGACCCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA
CTTTGCTGTTTCCTG

FIGURE 106

TTCCTGGGATGGATCCGCCCCATCNTCACATGCCCTGCCCCNTGGAGATTTACNCCTATGC
TGGCGAACAACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGGAATGTCCTGCGTGTC
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGACGGTCTGGCTATTTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAACTCTATGACCCTATGACCCAGTCAATGCCAGGTA
CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTG
CCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCACG

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

<subunit 1 of 1, 243 aa, 1 stop

<MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHELKFRGPFTDVVTTNLKLGNPTRNVCFKVKTTAPRRYCVRPNSGIID
AGASINVSVMQLQPFDYDPNEKSKHKFMVQSMFAPTDTSMEAVWKEAKPEDLMDSKLRVFE
LPAENDKPHDVEINKIISTTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE
NKQFKEEDGLRMRKTVQSNSPISALAPTGKEEGLSTRLLALVVLFFIVGVIIIGKIAL

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

FIGURE 111

TATTGTAAAGGCCATTTTAAACCATTTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCCTCGCCTGTTGGTGCTGGCCNTTGGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCCACTCCCGGCCCAGGCTGCTTT
CCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAAGCC
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTA CTGTCATAAGTGAGAGGCGTGTGTTGA
CTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCT
AAATTGTATTGGTTCATGTAGTGAAGTCAA CTGTTATTCAGAGATGTTTAATGCATATTTA
ACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTG

FIGURE 112

CCCTGGTGGTTTTGTTCTTTAATTCGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA
GNATGCACCNGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTTAAATT
ACCCATCCCTGNANACACATACACAGATACACANANACAAATNTAATGTAACGATNTTTTAG
AAAGTTAAAAATGTATAGTAAC

FIGURE 115

AAACCTTTAAAAGTTGAGGGGAAAAGAATGATCCTTTATTAATGACAAGGGAAACNTGNGT
AATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGC
TGGATTACCTCTCTTAAAAATGACACCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTN
GAGCCCAGCATGCTGGGGAGTGCGGTCTGCTCCACACAGTAGTCCCCANGTGGCCANTCCC
GGCCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGANTGATGA
ACAGAGTCAGAAGCCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA
GAGGCGTGTTGANTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCANTT
AAAGGGNCCAAGNTAAATTTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTCAGAGATG
TTTAATGCATATTTAANTTATTTAATGTATTTTCATNTCATGTTTTCTTATTGTCACAAGGGT
ACAGTTAATGCTGCGTGCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

FIGURE 116

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCACTCCCGGCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGC
TCCTTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGC
AGTGCTTTGTTCACTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAA
CTGTTATTCAGAGATGTTTAATGCATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTC
TTATTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTAT
TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 117

GCGAGCTCCGGGTGCTGTGGCCCGGCCTTGGCGGGGCGGCCTCCGGCTCAGGCTGGCTGAGA
GGCTCCCAGCTGCAGCGTCCCCGCCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
CTCGGGGACCAAACAAGCCTGGCAGGGTCTCACTTTGTTGCCAGGCTGGAGTTCAGTGCCA
TGATCATGGTTTACTGCAGCCTTGACCTCCTGGGTTCAAGCGATCCTGCTGAGTAGCTGGGA
CTACAGGACAAAATTAGAAGATCAAAATGGAAAATATGCTGCTTTGGTTGATATTTTTTCACC
CCTGGGTGGACCCTCATTTGATGGATCTGAAATGGAATGGGATTTTATGTGGCACTTGAGAAA
GGTACCCCGGATTGTGAGTGAAAGGACTTTCCATCTCACCAGCCCCGCATTTGAGGCAGATG
CTAAGATGATGGTAAATACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCAACTCCCAGC
CTTTCTGAATTGGAGGATTATCTTTCTATGAGACTGTCTTTGAGAATGGCACCCGAACCTT
AACCAGGGTGAAAGTTCAAGATTTGGTTCTTGAGCCGACTCAAAATATCACCACAAAGGGAG
TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCGACAGCAGGTTGAGCATCTTGGACAAA
AGGTTCTTAACCAATTTCCCTTTGAGCACAGCTGTGAAGCTTTCCACGGGCTGTAGTGGCAT
TCTCATTTCCTCAGCATGTTCTAACTGCTGCCACTGTGTTGATGAGAAAGGACTATG
TCAAAGGGAGTAAAAAGCTAAGGGTAGGGTTGTTGAAGATGAGGAATAAAAGTGGAGGCAAG
AAACGTCGAGGTTCTAAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGACCAAAGAGAGGGTAC
CAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAAAAATCTGGCCGGGGTC
AGAGGATTGCCGAAGGGAGGCCCTTCTTTGAGTGGACCCGGGTCAAGAATACCCACATTCCG
AAGGGCTGGGCACGAGGAGGCATGGGGGACGCTACCTTGACTATGACTATGCTCTTCTGGA
GCTGAAGCGTGCTCACAAAAGAAATACATGGAACCTTGAATCAGCCCAACGATCAAGAAAA
TGCCTGGTGGAATGATCCACTTCTCAGGATTTGATAACGATAGGGCTGATCAGTTGGTCTAT
CGGTTTTGAGTGTGTCCGACGAATCCAATGATCTCCTTTACCAATACTGCGATGCTGAGTC
GGGCTCCACCGGTTGCGGGGTCTATCTGCGTCTGAAAGATCCAGACAAAAGAATTGGAAGC
GCAAAATCATTGCGGTCTACTCAGGGCACCAGTGGGTGGATGTCCACGGGGTTGAGAAGGAC
TACAACGTTGCTGTTGCGATCACTCCCCTAAAATACGCCCAGATTTGCCTCTGGATTACGG
GAACGATGCCAATTGTGCTTACGGCTAAACAGAGACCTGAAACAGGGCGGTGTATCATCTAAA
TCACAGAGAAAACCAGCTCTGCTTACCGTAGTGAGATCACTTCATAGGTTATGCCTGGACTT
GAACTCTGTCAATAGCATTTCAACATTTTTCAAAATCAGGAGATTTTCGTCCATTTAAAAAA
TGTATAGGTGCAGATATTGAAACTAGGTGGGCACCTCAATGCCAAGTATATACTCTTCTTTA
CATGGTGATGAGTTTCATTTGTAGAAAAATTTGTTGCCTTCTTAAAAATTAGACACACTTT
AAACCTTCAAACAGGTATTATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC
TACTCTAAGAAGAATCTAATAGGATGCTGGTTGTGTATTAAATGTGAAATTGCATAGATAAA
GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTTGTACTA
CTCTGAGATGGATCCATTCAGCTCATGCCCTCAATGTTTATATTGTGTTATCTGTTGGGTCT
GGGACATTTAGTTTGTGTTTTTTGAAGAATTACAAATCAGAAGAAAAAGCAAGCATTATAAA
CAAACTAATAACTGTTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAA
TGGGAGAAATAGTTTGTCTATGAAATAAACCTAGTTTAGAAATAGGGAAGCTGAGACATTT
TAAGATCTCAAGTTTTTATTTAACTAATACTCAAAATATGGACTTTTCATGTATGCATAGGG
AAGACACTTCACAAATTATGAATGATCATGTGTTGAAAGCCACATTATTTTATGCTATACAT
TCTATGTATGAGGTGCTACATTTTTAGGACAAAGAATTCTGTAATCTTTTTCAAGAAAGAGT
CTTTTTCTCCTTGACAAAATCCAGCTTTTGTATGAGGACTATAGGGTGAATTCTCTGATTAG
TAATTTTAGATATGTCCTTTCCTAAAAATGAATAAAATTTATGAATATGA

FIGURE 118

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253

<subunit 1 of 1, 413 aa, 1 stop

<MW: 47070, pI: 9.92, NX(S/T): 3

MENMLLWLIFFTP^gWTLIDGSEMEWDFM^gWHLRKVPRIVSERTFHLTSPA^gFEADAKMMVNTVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTKGVSVRRKRQV
YGTDSRFSILDKRFLTNFPFSTAVKLSTGCSGILISPQHVL^gTA^gAHCVHDGKDYVKGSKKLRV
GLLKMRNKSGGKKRRGSKRSRREASGGDQREGTREHLQERAKGRRRKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGMGDATLDYDYALLELKRAHKKKYMELGISPTIKMPGGMIHFS
GFDNDRADQLVYR^gFCSVSD^gESNDLLYQYCD^gAESGSTGSGVYLRLKDPDKKNWKRKIIAVYSG
HQWVDVHGVQKDYNVAVRITPLKYAQICLWIHGNDANCAYG

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTTAGCACCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTAGTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGCATGGAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTC
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTG
GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCCCTCTACTCTCTGAGGCGCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC
TGCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCA
AGTGCCAAGCCAAAGGAACAACCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTTATGCCAAATTGTTGGCTGCGA
TCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATCGGATGATACT
GTGGTTGCACTTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAAGGTCTGATCACTT
ATATCTGGAAACCAAACCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAA
CTTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTCAATTGTCAAGATTGTAACCTCGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTC
CTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGC
CTTATGACCTCTACCATCCCCCTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCC
TCGTGTGGGGGGGGCATCCAGAGCCGGGCAGTTTTCTGTGTGGAGGAGGACATCCAGGGGCA
TGTCACCTCAGTGGAAGAGTGGAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCT
GCAACATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT
GGCCAGGGCCTCAGATACCGTGTGGTCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA
AACCCAAAGAGAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTGTCAGAGGAGCCCTCGTAAAGTTGTAAAAGCACAGACTGTTCTATA
TTTGAAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGA
ACTAAGTGTAATCATCTCACCAAAGCTTTTTGGCTCTCAAATTAAGATTGATTAGTTTCAA
AAAAAAAAA

FIGURE 120

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847

<subunit 1 of 1, 525 aa, 1 stop

<MW: 58416, pI: 6.62, NX(S/T): 1

MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTC GGGASYSLRRCLS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNPDNPCSLKCQ
AKGTTLVVELAPKVL DGT R CYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLV LKGP D HLYLETKTLQGTKGENSLSSTGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRET DFFPCS
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDP CPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKL PVEAKLPWFKQAQELEEGA AVSEEPS

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAACTCCCGTGGAGGGGCCGGTGGGCCCTCGGGCCTGAC
AGATGGCAGTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCG
GCCCCCGGTTTCGTGGGGCCAGGGTCCAGCGGCTGCGCAGAGGCGGGGACCCCGGCCTCAT
GCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGGCCTGGGCCGCGCCACGGCCGCCG
AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCGCCGAGGAG
GCGGCGGGTCAGCTCCGCCGCGAGCTCCGCCAGGCCGCGGAGTGCGGCCAGAGCCTGGCGT
CAGCGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG
CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG
ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA
TCTGGGGCACTTTCTACTACCAATCTTCTCCTTGGA CTCTCAAAGTTTCAGCTCCCAGCA
GGATTGTGGTAGTTTCTTCCAACTTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAAGCTTTTGTTATAGCCGGAGCAAACCTGGCTAACATTCTTTT
TACCAGGGAAGTAGCCCCGCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTTGCATCCTG
GTATTGTACGGACAAATCTGGGGAGGCACATACACATTCCACTGTTGGTCAAACCACTCTTC
AATTTGGTGTGATGGGCTTTTTTCAAACCTCAGTAGAAGGTGCCAGACTTCCATTTATTT
GGCCTCTTCACCTGAGGTAGAAGGAGTGTGAGGAAGATACTTTGGGGATTGTAAAGAGGAAG
AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAACCTCTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTTTATAAACTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAAGAATTTTG
ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTTACTGAAAAATTATTTT
TGGGATAAGAGAATTTTCAACAAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATA
GTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCAAGCATGGATGACATATTA
ATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGTT
TCATGGCCAAAGTGTTAACTAGTTTTTACTACAATGTTTGGTGTGTTGTGTGGAAATTATCTGC
CTGGTGTGTGCACACAAGTCTTACTTGGAATAAATTTACTGGTAC

FIGURE 122

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747

<subunit 1 of 1, 336 aa, 1 stop

<MW: 36865, pI: 9.15, NX(S/T): 2

MAVATAAAVLAAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEEAAGQLRREL RQA AECGPEPGVSGVGELIVRELDASLRSVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNH LGHFLLTNLLGLLKSSAPSR
IVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSR SKLANILFTRELARRLEG TNVT VNVLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK

Important features:

Signal peptide:

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACT
TGAAGAAAAAGAATTTTGTATTTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTTGGAG
TTACTGAAAAATTATTTTTGGGATAAGAGAATTTTCAAGCAAAGATGTTTTTAAATATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATTTGTAAAAATTATAACTGGGCA
AGCATGGATGACATATTAATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTT
TTCAAGTATCTTTGAGTTTCATGGCCAAAGTGTTAACTAGTTTTTACTACAATGTTTGGTGTT
TGTGTGGAAATTATCTGCCTGGCTT

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 124

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCCGGAGCCCAGCC
CTTTCCTAACCCAACCCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCACGGAC
CCCAGCGTTACCAATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCT
GCTCCTGGTAACTTGGGTTTTTACTCCTGTAACAACTGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTTTGAGGAAGCTTCCGATGTCATTAAGGAAGAATT
TCCAAATGAAAATCAAGTAGTGTGTTGCCAGAGTTGATTTGTGATCAGCACTCTGACATAGCCC
AGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTTCGTAATGGGATGATGATGAAG
AGAGAATACAGGGGTGAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAAG
TGACCCCATTCAGAAATTCGGGACTTAGCAGAAATCACCCTCTTGATCGCAGCAAAAGAA
ATATCATTGGATATTTTGGAGCAAAAGGACTCGGACAACTATAGAGTTTTTGAACGAGTAGCG
AATATTTTGCATGATGACTGTGCCTTCTTTCTGCATTTGGGGATGTTTCAAAACCGGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTTTGTGTTGACTTACAATTGGATTCAAGATAAATGTGTTCTCTCT
GTCCGAGAAATAACATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCAT
ACTCTTTCACATGAAAGAAGATACAGAAAGTTTAGAAATATTCAGAATGAAGTAGCTCGGC
AATTAATAAGTGAAAAAGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACAT
CCTCTTCTGCACATACAGAAAACCTCCAGCAGATTGTCTGTAAATCGCTATTGACAGCTTTAG
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAATTCCTGGAAAACCTCAAGCAATTCTG
TATTTGACTTACATTCTGGAAAACCTGCACAGAGAATTCCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTTTTAAAAACTTG
AAAAACAGTTTGTAAGCCTTTCAACAGCAGCATCAACCTACGTGGTGGAATAGTAAACCTA
TATTTTCATAATTCTATGTGTATTTTTATTTTGAATAAACAGAAAGAAATTTAAAAA
AAAAA

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46927, pI: 5.21, NX(S/T): 0

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKREYR
GQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKSDNYRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNI IYKPPGHSAPDMVYLGAMTNFDVTYNW IQDKCVPLVREI
TFENGEEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG
EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

FIGURE 126

ATTAAGGAAGAATTTCCAAATGAAAATCAAGTAGTNTTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTA

FIGURE 127

AGAGGCCTCTCTGGAAGTTGTCCCGGGTGTTGCGCCGCNGGAGCCCGGGTCGAGAGGACNAGG
TGCCGCTGCCTGGAGAATCCTCCGCTGCCGTGCGCTCCCGGAGCCCAGCCCTTTCCTAACCC
AACCCAACCTAGCCCNGTCCCAGCCGCCAGCGCCTGTCCCTGTCNCGGANCCCAGCGTNACC
ATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCTCCTGGTAAC
TTGGGTTTTTACTCCTGTAACAACTGAAATAACNNGTCTTGATACNNAGAATATAGATGAAA
TTTTAAACNATGCTGATGTGGCTTTAGTCAATTTTTATGCTGACTGGTGTCGTTTCAGTCAG
ATGTGGCATCCAATTTTTGAGGANGCTTCCGATGTCATTAAGGAAGAATTTCCAAATGAAAA
TCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA
TAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

1000 900 800 700 600 500 400 300 200 100 0

FIGURE 128

GCCCACGCGTCCGATGGCGTTCACGTTTCGCGGCCTTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCCGCGCTCATCTTCTTCGCCATTTGGCACATTATAGCATTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGTAATACCCTGAATCCCCTTGTACTCCCAGAGTACCTCAT
CCACGCTTTCTTCTGTGTCATGTTTCTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGATGAGTGGCCCAGGA
CTCTATGACCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAATTAGCTTTTTTATCTTCTAGCATTTTTTTTACTACCTATATGGCATGATCTATGTTT
TGGTGAGCTCTTAGAACAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACTTTAAAAAATGACTCCTTATTTTTTAAATGTTTCCACATTTTTTGCTTGTGGAAAGACTG
TTTTTCATATGTTATACTCAGATAAAGATTTTTAAATGGTATTACGTATAAATTAATATAAAAT
GATTACCTCTGGTGTTGACAGGTTTGAACTTGCACTTCTTAAGGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCCTAGTACATTGGAAGCTTTTGTATTATAGGAACTTGTA
GGGCTCATTTTGGTTTCATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTGAAAATGTATATCTGACTAGTGGGAACTTCATGGGTTTCTCATCTGTC
ATGTCGATGATTATATATGGATACATTTACAAAAATAAAAAGCGGGAATTTTCCCTTCGCTT
GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA
TATACTTGCTTTAATTCTTAAGCATAAGTAAACATGATATAAAAAATATATGCTGAATTACTT
GTGAAGAATGCATTTAAAGCTATTTTTAAATGTGTTTTTATTGTAAAGACATTACTTATTAAG
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTTAAGAATTTGCAGG
TACTACAGATTTTCAAACCTGAATGAGAGAAAATTGTATAACCATCCTGCTGTTCTTTAGT
GCAATACAATAAACTCTGAAATTAAGACTC

FIGURE 131

CGGACGCGTG GGGGAAACCCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAGATGGCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCGC
TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTGCGGGACCGCTTCGGCTGAAGCATTGAC
TCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACAC
CTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCC
ATTCGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCCAAAATGCACCTACTCTTTC
CTCTAACTCTGGTGAGGTCAATCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTTATCTTCAAGCCGATGACGGAAAAATAGTTATATTCAGTCTAAGCC
AGAAATCCAGTACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA
GCAAAATGTCCTATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTTCTTGAAGATGGA
GAAAGTGATGGCTTTTTAAGATGCCTCTCTCTTAACTCTGGGTGGATTTTAACTACAACCTCT
TGTCTCTCGGTGATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTTCTCTTGTGGTTGTTAGATCTAAACTGAAGATCATGAAGA
AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATTCTGAAATTTAAGCATTTTTCTTTT
AAAAGACAAGTGTAATAGACATCTAAAATCCACTCCTCATAGAGCTTTTAAAATGGTTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAAGTTACTCAAATCTGTG

FIGURE 132

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847

<subunit 1 of 1, 323 aa, 1 stop

<MW: 36223, pI: 5.06, NX(S/T): 1

MAAPKGSLWVRTQLGLPPLLLLLTMALAGGSGTASAEAFDSVLGDTASCHRAQCLTYPLHTYP
KEEELYACQRGCRFLFSICQFVDDGIDLNRTKLECESACTEAYSQSDEQYACHLGCQNQLPFA
ELRQEQLMSLMPKMHLLFPLTLVRSFWSDMMDSAQSFITSSWTFYLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYLMRNSQAHRNFLEDGESDGFRLCLSLNSGWILTTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

FIGURE 133

TTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTTCGC
TGAAGTGAAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAA
CTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGC

133

FIGURE 134

CACACTGGCCGGATCTTTTAGAGTCCTTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAACAAGATGGCGGCGCCGAAGGGAGCCTTTGGGTGAGGAC
CCAAC TGGGGCTCCCGCCGCTGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTGCGGGACCG
CTTCGGCTGAAGCATTGTGACTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAG
TTGACCTACCCCTTGACACCTACCCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTTCAATTTGTCAGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTT
GGTTGCCAGAATCAGCTGCCATTCGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCC
AAAAATGCACCTACTCTTTCCTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACT
CCGC

FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCCGGAGGT
GGGGCGCCGCTGGGGCCGGCCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCCGCGACCGAGC
GTGCGGACTGGCCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCAATGGGCGCGGCTG
GGGATTCTTGTGTTGGCCTCCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGGAGAGGAGC
AGCCCCCGGAGACAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTTGGATGATTGT
ACCTGTGATGTTGAAACCATTGATAGATTTAATAACTACAGGCTTTTCCCAAGACTACAAAA
ACTTCTTGAAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGTCTTTCT
GGAATGACATCAGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAA
GTTCTGATGGAATTAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA
AGGCTGTTCTTCAGTGGACCAAGCATGATGATTCTTCAGATAACTTCTGTGAAGCTGATGAC
ATTCAGTCCCCTGAAGCTGAATATGTAGATTTGCTTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACCAGATGCTTGGAAAATATGGAATGTCATCTACGAAGAAAACCTGTTTTAAGCCAC
AGACAATTAAAAGACCTTTAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGAACACT
TTTTACAGTTGGCTAGAAGGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAATGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAG
AAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTTGATGGAATTTTGACTGAA
GGAGAAGGTCCAAGAAGGCTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTT
ATCCAAAGTGTTACCATTCTTCGAGCGCCAGATTTTCAACTCTTTACTGGAAATAAAATTC
AGGATGAGGAAAACAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTTCTTTG
CATTTTGATGAGAATTCATTTTTTGCTGGGGATAAAAAAGAAGCACACAACTAAAGGAGGA
CTTTCGACTGCATTTTAGAAATATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTC
GTCTGTGGGGAAAGCTTCAGACTCAGGGTTTGGGCACTGCTCTGAAGATCTTATTTTCTGAG
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCCATCTAACCAGACAAGA
AATAGTATCATTATTC AACGCATTTGGAAGAATTTCTACAAGTGTGAAAGAATTAGAAAAC
TCAGGAACTTGTTACAGAATATTCATTAAAGAAAACAAGCTGATATGTGCCTGTTTCTGGAC
AATGGAGGCGAAAGAGTGGAATTTCAATCAAAGGCATAATAGCAATGACAGTCTTAAGCCAA
ACATTTTATATAAAGTTGCTTTTGTAAGGAGAATTATATTGTTTTAAGTAAACACATTTTT
AAAAATTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAAGTAATACTTTAATAATGTG
GTACAAATTTTAAAGTTAATATTGAATAAAAGGAGGATTATCAAATTAATAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 136

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974

<subunit 1 of 1, 468 aa, 1 stop

<MW: 54393, pI: 5.63, NX(S/T): 2

MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAAQRCFCQVSGYLDDCTCDVETIDRFNNYRLF
PRLQKLLESDFRYKVNLRPCPFWNDISQCGRRDCAVKPCQSDEVPDGIKSASYKYSEEA
NNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLLNP
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLI
ELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFCRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAACTGTTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAAATTTNAACAGCGATTTGATGGAATTTTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTTACCATT
CTTNGAGCGCCCAGATTTTCAACTNTTTACTGGAAATAAAATTCAGGATGAGGNAAACAAAA
TGTTACTTTTGGAAATACTTCATGAAATCAAGTCATTTCTTTGCATTTTGATGAGAATTCA
TTTTTTTGCTG

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039

><subunit 1 of 1, 124 aa, 1 stop

><MW: 13352, pI: 5.99, NX(S/T): 1

MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW

DGPSGDRRGDVYRCFVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETGDGGFMVS

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

FIGURE 141

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTGAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AATGCAGACTTTTACAATGGTTCTAGAAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA
AACAGTGTACTATTCTGTGCAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATAACAACCTTCGTGTGAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
CATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCCTTTGTTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTTCGTCTGGAATAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG
TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCAGAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
CTCTAGGTTTGCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAAACTGGTGACACTCTACAGTCTGACTGATTCAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGGAGAGCCCACTTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCATTCAAGCCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACTACTGTAAAGTGTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTTC
TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
AAAAAAAAA

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033

<subunit 1 of 1, 311 aa, 1 stop

<MW: 35076, pI: 5.04, NX(S/T): 2

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLIMWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTS
SILKHPFNRNSTILTRPGMEITKDGFHLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPVLALFAFVGFMILIV
VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

FIGURE 143

TCCTGCTGATGCACATCTGGGTTTGGCAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTGAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTGCGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353

<subunit 1 of 1, 328 aa, 1 stop

<MW: 36238, pI: 9.90, NX(S/T): 3

MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLFLGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL
FVNVASTSNPFLSRLLNRDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSLTPPCSE
TVTWILIDRALNITSLQMHSLRLLSQNPSPQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVDGVPHGR

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 146

GGCGCCTGGTTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTCGCCGCCAGCCTCCGCCGCCGAGCCTC
GTTCTGTTCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGGGGCGCCACCCTGGCAGACTAACGAA
GCAGCTCCCTTCCACCCCACTGCAGGTCTAATTTTGGACGCTTTGCCTGCCATTTCTTCCAGGTTGAGGGAGC
CGCAGAGGCGGAGGCTCGCGTATTCTGTCAGTCAGCACCCACGTCGCCCCCGGACGCTCGGTGCTCAGGCCCTTC
GCGAGCGGGGCTCTCCGTCTGCGGTCCCTTGTGAAGGCTCTGGGCGGCTGCAGAGGCCGGCCGTCCGGTTTGGCT
CACCTCTCCAGGAACTTCACACTGGAGAGCCAAAAGGAGTGGAAGAGCCTGTCTTGGAGATTTTCTGGGGAA
ATCCTGAGGTCAATTCATTATGAAGTGTACCGCGCGGAGTGGCTCAGAGTAACACAGTGCTGTTTCATGGCTAGA
GCAATTCAGCCATGGTGGTTCCCAATGCCACTTTATTGGAGAACTTTTGGAAAAATACATGGATGAGGATGGT
GAGTGGTGGATAGCCAAACAACGAGGGAAAAGGGCCATCACAGACAATGACATGCAGAGTATTTTGGACCTTCAT
AATAAATTACGAAGTCAGGTGTATCCAACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAAGA
TCTGCAGAACTCTGGGCTGAAAGTTGCTTGTGGGAACATGGACCTGCAAGCTTGCTTCCATCAATTGGACAGAAT
TTGGGAGCACACTGGGGAAGATATAGGCCCCCGACGTTTTCATGTACAATCGTGGTATGATGAAGTGAAAGACTTT
AGCTACCCATATGAACATGAATGCAACCCATATTGTCCATTAGGTGTTCTGGCCCTGTATGTACACATTATACA
CAGGTCTGTGGGCAACTAGTAACAGAATCGGTTGTGCCATTAATTTGTGTACATAACATGAACATCTGGGGCGAG
ATATGGCCCAAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAAGGGAAGTGGTGGGGCCATGCCCTTTACAAA
CATGGGCGGCCCTGTTCTGCTTGCCCACTAGTTTTTGGAGGGGGCTGTAGAGAAAATCTGTGCTACAAAGAAGGG
TCAGACAGGTATTTATCCCTTCGAGAAGAGGAAAACAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC
CATGTCCGGACAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGCGCACAGCAAATGTCCCAAATTTGTTTCT
TGTGAAGTAAGATTAAGAGATCAGTGCAAAGGAACAACCTGCAATAGGTACGAATGTCTGTGGCTGTTTGGAT
AGTAAAGCTAAAGTTATTTGGCAGTGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTCATTATGGT
ATAATAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTTCATCAAGTCCAATAGA
AATGGTATTTCAAACAATTGGCAAATATCAGTCTGCTAATTCCTTACAGTCTCTAAAGTAACAGTTTCAGGCTGTG
ACTTGTGAAACAACCTGTGGAACAGCTCTGTCCATTTTCATAAGCCTGCTTCACATTGCCCAAGAGTATACTGTCT
CGTAACTGTATGCAAGCAAATCCACATTATGCTCGTGTAAATGGAACTCGAGTTTATTCTGATCTGTCCAGTATC
TGCAGAGCAGCAGTACATGCTGGAGTGGTTCGAAATCACGGTGGTTATGTTGATGTAATGCCTGTGGACAAAAGA
AAGACCTACATTGCTTCTTTTCAGAAATGGAATCTTCTCAGAAAGTTTACAGAATCCTCCAGGAGGAAAGGCATT
AGAGTGTGCTGTTGTGTGAAGTGAATACTTGAAGAGGACCATAAAGACTATTCCAAATGCAATATTTCTGA
ATTTTGTATAAAACCTGTAACATTACTGTACAGAGTACATCAACTATTTTCAGCCCAAAGGTGCCAAATGCATA
TAAATCTTGATAAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTGGGAAAAGTAATGAAAATATAATGG
TTTTAGAAATCCTGTGTTAAATATTGCTATATTTCTTAGCAGTTATTTCTACAGTTAATTACATAGTCATGATT
GTTCTACGTTTTCATATATTATATGGTGTCTTGTATATGCCACTAATAAAATGAATCTAAACATTGAATGTGAATG
GCCCTCAGAAAATCATCTAGTGCATTTAAAAATAATCGACTCTAAAACCTGAAAGAAACCTTATCACATTTTCCCC
AGTTCAATGCTATGCCATTACCAATCCAAATAATCTCAAATAATTTTCCACTTAATAACTGTAAAGTTTTTTTC
TGTTAATTTAGGCATATAGAATATTAAATTTCTGATATTGCACTTCTTATTTTATATAAAATAATCCTTTAATATC
CAAATGAATCTGTTAAATGTTTGATTCTTGGGAATGGCCTTAAAAATAAATGTAATAAAGTCAGAGTGGTGGT
ATGAAAACATTCTAGTGATCATGTAGTAAATGTAGGTTAAGCATGGACAGCCAGAGCTTTCTATGTACTGTTA
AAATTGAGGTCACATATTTTCTTTTGTATCCTGGCAAATACTCCTGCAGGCCAGGAAGTATAATAGCAAAAAGTT
GAACAAAGATGAACATAATGTATTACATTACCATTGCCACTGATTTTTTTTAAATGGTAAATGACCTTGTATATAA
ATATTGCCATATCATGGTACCTATAATGGTGATATATTTGTTTCTATGAAAAATGTATTGTGCTTTGATACTAAA
AATCTGTAAATGTTAGTTTTTGGTAATTTTTTTCTGCTGGTGGATTACATATTAAATTTTTTCTGCTGGTGGGA
TAAACATTAAATTAATCATGTTTCAAAAAAAAAAAAA

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417

<subunit 1 of 1, 500 aa, 1 stop

<MW: 56888, pI: 8.53, NX(S/T): 2

MKCTAREWLRVTTVLFMARAI PAMVVPNATLLEKLLLEKYMDEDEGEWWIAKQRGKRAITDNDM
QSILDLHNKLR SQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW
GRYRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQV VWATSNRIGCAINLC
HNMNIWGQIWP KAVYLV CNYS PKGNWWGHAPYKHGRPCSACPPSF GGGCRENL CYKEGSDRY
YPPREEETNEIERQQSQVHDTHVRTRSDSSRNEVIS AQQMSQIVSCEVRLRDQCKGTT CNR
YEC PAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKHYFIKSNRNGI
QTIGKYQSANSFTVSKVTVQAVTCETTVEQLCPFHKPASHCPRVYCPRNCMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSESLQNPPGGKA FRV
FAVV

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGGAAGGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGCCGCGCTCCGACGGGCCAGCGCCCTCCCCATGTCCCTGCTCCCACGCCC
CGCCCCCTCCGGTTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGT
ACACCGCGCGTGTGGACGGGTCCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTAC
AGCGACGTGAAGAAGCTGGAAATGAAGCCAAAGTACCCGCACTGCGAGGAGAAGATGGTTAT
CATCACCACCAAGAGCGTGTCCAGGTACCGAGGTGAGGAGCACTGCCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAATAGGGGTGAAAAACCTCAGAAGGGAAAACTCCAAACCAGTTGGGAGACTTGTGCAAAGGA
CTTTGCAGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCCTTTC
TTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTTACCAACGGTCAG
TTTTTACATTTTATAGCTGCGTGCGAAAGGCTTCAGATGGGAGACCCATCTCTCTTGTGCT
CCAGACTTCATCACAGGCTGCTTTTTATCAAAAAGGGGAAAACTCATGCCTTTCCTTTTTAA
AAAATGCTTTTTTTGTATTTGTCCATACGTCACTATACATCTGAGCTTTATAAGCGCCCGGGA
GGAACAATGAGCTTGGTGGACACATTTCAATTGCAGTGTTGCTCCATTCCTAGCTTGGGAAGC
TTCCGCTTAGAGGTCCTGGCGCCTCGGCACAGCTGCCACGGGCTCTCCTGGGCTTATGGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCCCTGTAGCCGGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTCTGAGGAACTCAAGTTTGGTTGCCAGAAAAATGTGCTTCATTCCCCCT
GGTTAATTTTTTACACACCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGGTCTTTCCCAACCTGAGGATTTCTGAAAGGTTACAGGTTCAATA
TTTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTCAGCAAAAACCTTAGGAGAAAACT
TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTTCAAAGCATGTTTCTTTCCCTCACCACAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTCCCCATGTAATTCTTCAATGTTAAACAGTGCAGTCCTCTTTGAAAGCTAAGAT
GACCATGCGCCCTTTCCTCTGTACATATACCCTTAAGAACGCCCCCTCCACACACTGCCCCC
CAGTATATGCCGCATTGTACTGCTGTGTTATATGCTATGTACATGTCAGAAACCATTAGCAT
TGCATGCAGGTTTCATATTCTTTCTAAGATGGAAAGTAATAAAATATATTTGAAATGTAAAA
AAAAAAAAAAAA

FIGURE 149

MSLLPRRAPPVSMRLAAALLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH
CEEKMVIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNWNEKRRVYEE

Signal sequence:

amino acids 1-34

Downloaded from www.jstor.org

FIGURE 150

GCCCCAGGGACTGCTATGGCTTCCTTTTGTGTTTACCCCCGGTCTGCGTCAATGTTAAACTCCAATGTCTCTCTGTG
GTTAACTGCTCTTGCCATCAAGTTACCCCTCATTGACAGCCAAGCACAGTATCCAGTTGTCAACACAAATTATGG
CAAAATCCGGGGCCTAAGAACACCGTTACCCAATGAGATCTTGGGTCCAGTGGAGCAGTACTTAGGGGTCCCCA
TGCCCTACCCCCCACTGGAGAGAGGCGGTTTTCAGCCCCAGAACCCCCGTCCTCTGGACTGGCATCCGAAATAC
TACTCAGTTTGCTGCTGTGTGCCCCAGCACCTGGATGAGAGATCCTTACTGCATGACATGCTGCCCATCTGGTT
TACCGCCAATTTGGATACTTTGATGACCTATGTTCAAGATCAAAATGAAGACTGCCTTTACTTTAAACATCTACGT
GCCCACGGGAAGATGGAGCCAACACAAAGAAAAACGCAGATGATATAACGAGTAATGACCGTGGTGAAGACGAAGA
TATTCATGATCAGAACAGTAAGAAGCCCGTCATGGTCTATATCCATGGGGGATCTTACATGGAGGGCACCGGCAA
CATGATTTGACGGCAGCATTTTGGCAAGCTACGGAAACGTCATCGTGATCACCATTAACTACCGTCTGGGAATACT
AGGGTTTTTAAGTACCGGTGACCAGGCAGCAAAAGGCAACTATGGGCTCCTGGATCAGATTCAAGCACTGCGGTG
GATTGAGGAGAATGTGGAGCCTTTGGCGGGGACCCCAAGAGAGTGACCATCTTTGGCTCGGGGGCTGGGGCCTC
CTGTGTGACGCTGTGTGACCCCTGTCCCACTACTCAGAAGGTCTCTTCCAGAAGGCCATCATTCAGAGCGGCACCGC
CCTGTCCAGCTGGGCAGTGAACCTACCAGCCGGCCAAGTACACTCGGATATTGGCAGACAAGGTCGGCTGCAACAT
GCTGGACACCACGGACATGGTAGAATGCCTGCGGAACAAGAACTACAAGGAGCTCATCCAGCAGACCATCACCCC
GGCCACCTACCACATAGCCTTCGGGCCGGTGATCGACGGCGACGTCATCCAGACGACCCCCAGATCCTGATGGA
GCAAGGCGAGTTCTTCAACTACGACATCATGCTGGGCGTCAACCAAGGGGAAGGCCGTAAGTTTCGTGGACGGCAT
CGTGGATAACGAGGACGGTGTGACGCCCCAACGACTTTGACTTCTCCGTGTCCAACCTTCGTGGACAACCTTTACGG
CTACCCCTGAAGGGAAAGACACTTTGCGGGAGACTATCAAGTTCATGTACACAGACTGGGCCGATAAGGAAAACCC
GGAGACGCGCGGAAAAACCTGGTGGCTCTCTTTACTGACCACCAGTGGGTGGCCCCCGCGTGGCCGCCGACCT
GCACGCGCAGTACGGCTCCCCACCTACTTCTATGCCTTCTATCATCACTGCCAAAGCGAAATGAAGCCCAGCTG
GGCAGATTCGGCCCATGGTGATGAGGTCCCTATGTCTTCGGCATCCCCATGATCGGTCCCACCGAGCTCTTCAG
TTGTAACCTTTTCCAAGAACGACGTCATGCTCAGCGCCGTGGTTCATGACCTACTGGACGAACCTTCGCCAAAACTGG
TGATCCAAATCAACCAGTTCCTCAGGATACCAAGTTCATTACACAAAACCCCAACCGCTTTGAAGAAGTGGCCTG
GTCCAAGTATAATCCCAAAGACCAGCTCTATCTGCATATTGGCTTGAAACCCAGAGTGAGAGATCACTACCGGGC
AACGAAAGTGGCTTTCTGGTTGGAACCTCGTTCCTCATTTGCACAACCTTGAACGAGATATTCCAGTATGTTTCAAC
AACCACAAAGGTTCTTCCACCAGACATGACATCATTTCCCTATGGCACCCGGCGATCTCCCGCCAAGATATGGCC
AACCACCAAACGCCCAGCAATCACTCTGCCAACAATCCCAACACTCTAAGGACCCCTCACAAAACAGGGCCTGA
GGACACAACCTGTCTCATTGAAACCAAACGAGATTATTCCACCGAATTAAGTGTACCATTGCCGTGCGGGCGTC
GCTCCTCTTCTCAACATCTTAGCTTTTGGCGCGCTGTACTACAAAAGGACAAGAGGCGCCATGAGACTCACAG
GCGCCCCAGTCCCAGAGAAACACCACAAATGATATCGCTCACATCCAGAACGAAGAGATCATGTCTCTGCAGAT
GAAGCAGCTGGAACACGATCACGAGTGTGAGTCGCTGCAGGCACACGACACACTGAGGCTCACCTGCCCGCCAGA
CTACACCCTCACGCTGCGCCGGTCCGACAGATGACATCCACTTATGACGCCAAACACCATCACCATGATTCACAA
CACACTGACGGGGATGCAGCCTTTGACACACTTTTAACACCTTCAGTGGAGGACAAAACAGTACAAATTTACCCCA
CGGACATTCACCACTAGAGTATAGCTTTGCCCCATTTCCCTTCTATCCCTCTGCCCTACCCGCTCAGCAACAT
AGAAGAGGGAAGGAAAGAGAGAAGGAAAGAGAGAGAGAAAGAAAGTCTCCAGACCAGGAATGTTTTTGTCCCACT
GACTTAAGACAAAAATGCAAAAAGGCAGTCATCCCATCCCGGCAGACCCTTATCGTTGGTGTTTTCCAGTATTAC
AAGATCAACTTCTGACCCTGTGAAATGTGAGAAGTACACATTTCTGTTAAAATAACTGCTTTAAGATCTCTACCA
CTCCAATCAATGTTTAGTGTGATAGGACATCACCATTTCAAGGCCCGGGTGTTCACACGTCATGGAAGCAGCT
GACACTTCTGAAACTCAGCCAAGGACACTTGATATTTTTTAATTACAATGGAAGTTTAAACATTTCTTCTGTGC
CACACAATGGATGGCTCTCCTTAAGTGAAGAAAGAGTCAATGAGATTTTGGCCAGCACATGGAGCTGTAATCCAG
AGAGAAGGAAACGTAGAAATTTATTATTAAGAAGTGGACTGTGCAGCGAAATCTGTACGGTTCTGTGCAAAGAG
GTGTTTTGCCAGCCTGAACATATTTAAGAGACTTTGT

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVNTNYGKIRGLRTPLPNEILGPVEQYLGVPYASP
PTGERRFQPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMLPIWFTANLDTLMTYVQDQN
EDCLYLNIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTLSHYSEGLFQKAI IQSGTALSSWAVNYQPAKYTRILADKVGCMNL
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGEFLNYDIMLG
VNOGEGLEKFDVGIVDNEDGVT PND FDFSVSNFVDNLYGYPEGKDTLRET IKFMYTDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKPSWADSAHGDEV
PYVFGIPMIGPTELFSCNFSKNDVMLS AVVMTYWTNFAKTGDPNQVPQDTKFIHTKPNR
FEEVAVSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTK
VPPDMTSPFYGTTRSPAKIWPTTKRPAITPANNPKHSDPHKTGPEDTTVLIETKRDY
STELSVTIAVGASLLFLNILAFAALYYKKDKRRHETHRRPSPQRNTTNDIAHIQNEE
IMSLQMKQLEHDHECESLQAHDTLRLTCPPDYTLTLRRSPDDIPLMTPNTITMI
PNTLTGMQPLHTFNTFSGGQNSTNLPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

FIGURE 152

GGGAAAGATGCGGGCGACTCTGGGACCCCTTGGGTCTGGCAGCAGTGGCGGCGATGTTTGT
CGGCTCGGGATGGGTCCAGGATGTTACTCCTTCTTCTTTTGTGGGGTCTGGGCAGGGGCCA
CAGCAAGTCGGGGCGGGTCAAACGTTTCGAGTACTTGAAACGGGAGCACTCGCTGTGGAAGCC
CTACCAGGGTGTGGGCACAGGCAGTTCCCTCACTGTGGAATCTGATGGGCAATGCCATGGTGA
TGACCCAGTATATCCGCCTTACCCAGATATGCAAAGTAAACAGGGTGCCTTGTGGAACCGG
GTGCCATGTTTCTGAGAGACTGGGAGTTGCAGGTGCACTTCAAAATCCATGGACAAGGAAA
GAAGAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGC
CTGTGTTTGGAAACATGGACAAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAAT
GAGGAGAAGCAGCAAGAGCGGGTATTCCCTACATCTCAGCCATGGTGAACAACGGCTCCCT
CAGCTATGATCATGAGCGGGATGGGCGGCCTACAGAGCTGGGAGGCTGCACAGCCATTGTCC
GCAATCTTCATTACGACACCTTCTGGTGATTTCGCTACGTCAAGAGGCATTTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCCGGAGTCCGCCTGCC
CCGCGGCTACTACTTCGGCACCTCCTCCATCACTGGGGATCTCTCAGATAATCATGATGTCA
TTTCTTTGAAGTTGTTTGAAGTGCAGTGGAGAGAACCCAGAGAGGAAAAGCTCCATCGA
GATGTGTTCTTGCCCTCAGTGGACAATATGAAGCTGCCTGAGATGACAGCTCCACTGCCGCC
CCTGAGTGGCCTGGCCCTCTTCTCATCGTCTTTTCTCCCTGGTGTTTTCTGTATTTGCCA
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTACTGA
GCCCTCCTGCTGCCACCACTTTTGTGACTGTACCCATGAGGTATGGAAGGAGCAGGCACTG
GCCTGAGCATGCAGCCTGGAGAGTGTCTTGTCTCTAGCAGCTGGTTGGGGACTATATTCTG
TCACTGGAGTTTTGAATGCAGGGACCCCGCATTCCCATGGTTGTGCATGGGGACATCTAACT
CTGGTCTGGGAAGCCACCCACCCAGGGCAATGCTGCTGTGATGTGCCTTCCCTGCAGTCC
TTCCATGTGGGAGCAGAGGTGTGAAGAGAATTTACGTGGTTGTGATGCCAAAATCACAGAAC
AGAATTTCATAGCCAGGCTGCCGTGTTGTTGACTCAGAAGGCCCTTCTACTTCAGTTTTG
AATCCACAAAGAATTAAAACTGGTAACACCACAGGCTTTCTGACCATCCATTTCGTTGGGTT
TTGCATTTGACCCAACCTCTGCCTACCTGAGGAGCTTTCTTTGGAAACCAGGATGGAACT
TCTTCCCTGCCTTACCTTCCTTTCACTCCAATTCATTGTCTCTCTGTGTGCAACCTGAGCTG
GGAAAGGCATTTGGATGCCTCTCTGTTGGGGCCTGGGGCTGCAGAACACACCTGCGTTTCAC
TGGCCTTCATTAGGTGGCCCTAGGGAGATGGCTTTCTGCTTTGGATCACTGTTCCCTAGCAT
GGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCAATCAAGTCTCTTCAGGCCCTCAGTG
AAGTTTGGCTAAAGGTTGGTGTAAAAATCAAGAGAAGCCTGGAAGACATCATGGATGCCATG
GATTAGCTGTGCAACTGACCAGCTCCAGGTTTGATCAAACCAAAGCAACATTTGTATGTG
GTCTGACCATGTGGAGATGTTTCTGGACTTGCTAGAGCCTGCTTAGCTGCATGTTTTGTAGT
TACGATTTTTGGAATCCCACTTTGAGTGCTGAAAGTGTAAGGAAGCTTTCTTCTTACACCTT
GGGCTTGGATATTGCCAGAGAAGAAATTTGGCTTTTTTTTTCTTAATGGACAAGAGACAGT
TGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCCTCATCATCTGTGCCTGGAAGAGTT
CACTGTCAATTGAGCAGCACAGCCTGAGTGCTGGCCTCTGTCAACCCTTATTCCACTGCCTTA
TTTGACAAGGGGTTACATGCTGCTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAG
AGTCTCCTTGGAGGGCCTGGAACCTCTGAGTCTCCTATGAACCTCTGTAGCCTAAATGAAAT
TCTTAAATCACCGATGGAACCAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCCATGG

FIGURE 153

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911

><subunit 1 of 1, 348 aa, 1 stop

><MW: 39711, pI: 8.70, NX(S/T): 1

MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLLGSGQGPQQVGAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQ GALWNRVPCFLRDWELQVHF KIHGQGKKN
LHGDGLAIWYTKDRMQPGPVFGNMDKFVGLGVFVD TYPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRN LH YDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKL FELTVERTPEEEKLHRDVFLPSVDNMKLPEMTAPLPPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRKRFY

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

FIGURE 154

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGGCCGGCCTGGGACCATGGGCGTGAGTGCAATCTACGGATCAGTCT
CTGATGGTGGGTCGTTAACCTCAGTGGGGACTCCAAGATTTCCATGAAGAAAATCAGTTGTCTTTCATTCAAGAAAT
TGGGGTCTGGCTCAGAATTCCTGCAGCTGGTGAAAATCTGTTTTCTAGAAGAGGTTTAATTAATGCCTGCAGTCT
GACATGTTCCCGATTGTGAGGTGAAAACCATGAAGAGAAAATAGAATACTTAATAATGCTTTTCCGCAACCGCTTCT
TGCTGCTGCTGGCCCTGGCTGCGCTGCTGGCCTTTGTGAGCCTCAGCCTGCAGTTCCTCCACCTGATCCCGGTGT
CGACTCCTAAGAAATGGAATGAGTAGCAAGAGTCGAAAGAGAATCATGCCCGACCCTGTGACGGAGCCCCCTGTGA
CAGACCCCGTTTATGAAGCTCTTTGTACTGCAACATCCCCAGTGTGGCCGAGCGCAGCATGGAAGGTCATGCC
CGCATCATTTTAAGCTGGTCTCAGTGCATGTGTTTCATTGCCCCGAGAGACAGGTACCCACTGTATGTTCATTTCCCA
AAACAAAGCGACCCAGAAAATTGACTGCACTCTGGTGGCTAACAGGAAACCGTATCACCCAAAACCTGGAAGCTTTCA
TTAGTCACATGTCAAAGGATCCGGAGCCTCTTTGAAAAGCCCCCTTGAACCTCTTGCCTCTTTACCCAAATCACC
CATTGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAGCATTTCAGAACGGTCAGCTGCTGAGGGATA
TCTATCTAAAGAAACACAACTCCTGCCAATGATTGGTCTGCAGACCAGCTCTATTTAGAGACCCTGGGAAAA
GCCGGACCCTACAAAGTGGGCTGGCCTTGCTTTATGGCTTTCTCCCAGATTTTGACTGGAAGAAGATTTATTTCA
GGCACCAGCCAAGTGCGCTGTTCTGCTCTGGAAGCTGCTATTGCCCCGTAAGAAACAGTATCTGGAAAAGGAGC
AGCGTCGTAGTACCTCCTACGTTTGAAAAACAGCCAGCTGGAGAAGACCTACGGGGAGATGGCCAAAGATCGTGG
ATGTCCCCACCAAGCAGCTTAGAGCTGCCAACCCCATAGACTCCATGCTCTGCCACTTCTGCCACAATGTCAGCT
TTCCCTGTACCAGAAATGGCTGTGTTGACATGGAGCACTTCAAGGTAATTAAGACCCATCAGATCGAGGATGAAA
GGGAAAGACGGGAGAAGAAATTTGACTTCGGGTATTCTCTCCTGGGTGCCACCCCATCCTGAACCAAACCATCG
GCCGGATGCAGCGTGCCACCGAGGGCAGGAAAGAGAGCTCTTTGCCCTCTACTCTGCTCATGATGTCACCTCTGT
CACCAGTTCTCAGTGCTTGGGCCCTTTGAGAAGCCAGGTTCCCAAGGTTTGAGCCAGGTTGATCTTTGAGCTTT
GGCAAGACAGAGAAAAGCCAGTGAACATTCGTCGGGATTCTTTACAATGGCGTCGATGTCACATTCACACCT
CTTTCTGCCAAGACCACCACAAGCGTTCTCCCAAGCCCATGTGCCCGCTTGAAAACCTTGGTCCGCTTTGTGAAA
GGGACATGTTTGTAGCCCTGGGTGGCAGTGGTACAAATTATTATGATGCATGTCACAGGGAAGGATTC~~TAAA~~AGG
TATGCAGTACAGCAGTATAGAATCCATGCCAATACAGAGCATAGGGAAAGGTCCACTTCTAGTTTTGTCTGTTAC
TAAGGGTAGAAGATTATTGCTTTTTTAAAGGCTAAATATTGTTTGTGGGAACCACAGATGGTTGGGGTTGAACAGT
AAGCACATTGCTGCAATGTGGTACGTGAATTGCTTGGTACAAAATGGCCAGTTCACAGAGGAATAGAAGGTACTT
TATCATAGCCAGACTTCGCTTAGAATGCCAGAATAATATAGTTCAAGACCTGAAGTTGCCAATCCAAGTTTGCAC
TCTTCTGGCCTGCCCATGTTACTATGTGATGGAACCAGCACACCTCAACCAAATTTTTTTAATCTTAGACATT
TTTACCTTGTCTTGTAAAGAATTTCTTGAAGTGATTATCTAAAAATAAAGGTTGGCAAACCTTTTTCTGTAAAGG
GCCAGATTGTAAATATTTAGACTGTGTGGACCAAAGGCCACATACAGTCTCTGTGCATAACTACTCAACTCTGT
TTCTGAAGCAGGAAAGCCACCACAGACAGTACATAAAGGAATATGTGTAGCTGGGTTCCAGGCCAGACAAAACA
GATGGTGACCAGACTTGGCCCCCTGGGCTGTAGTTTGTGACCCCTCATCTAAAAAATAGGCTATACTACAATTGC
ACTTCCAGCACTTTGAGAACGAGTTGAATACCAAGAATTATTCAATGGTTCTCCAGTAACCTCTGCTAGAAACA
CAGAATTTGGTCTGTATCTGACACTAGAACAAAACCTTGAGGGTAAATAAACATTGAATTAGAATGAATCATAGAA
AACTGATTAGAAGAATACTTGATGTTTATGATGATTGTGGTACAAGATAGTTTTAAGTATGTTCTAAATATTTGT
CTGCTGTAGTCTATTTGCTGTATATGCTGAAATTTTTGTATGCCATTTAGTATTTTTATAGTTTAGGAAAATATT
TTCTAAGACCAGTTTTAGATGACTCTTATTCCTGTAGTAATATTCAATTTGCTGTACCTGCTTGGTGGTTAGAAG
GAGGCTAGAAGATGAATTCAGGCACCTTCTTCCAATAAACTAATTATGGCTCATTCCCTTTGACAAGCTGTAGA
ACTGGATTCATTTTTAAACCATTTTCATCAGTTTCAAATGGTAAATTCTGATTGATTTTTTAAATGCGTTTTTGGG
AGAATTTGCTATTAGGTAGTTTACAGATCTTTATAAGGTGTTTTATATATTAGAAGCAATTATAATTACATCTG
TGATTTCTGAACTAATGGTGCTAATTCAGAGAAATGGAAAGTGAAAGTGAGATTCTCTGTTGTATCGGCATTCC
AACTTTTTCTCTTTGTTTTTGTCCAGTGTGCTTTGAATATGTCTGTTTCTATAAATAAATTTTTTAAAGATAA

FIGURE 155

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329

><subunit 1 of 1, 480 aa, 1 stop

><MW: 55240, pI: 9.30, NX(S/T): 2

MLFRNRFLLLLALAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDVTEPPVTDPVY
EALLYCNIPSHAERSMEGHAPHHFKLVSVHVFIRHGDRYPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKSGSGASFESPLNSLPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLK
KHKLLLPNDWSADQLYLETTGKSRTLQSGLLLYGFLPDFDWKKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSMMLCHFCHNVSFPC
TRNGCVDMEHFKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF
ALYSAHDVTLSPVLSALGLSEARFPRFAARLIFELWQDREKPSSEHSVRILYNGVDVTFHTSF
CQDHHKRSPKPMCPLENLVRFVKRDMFVALGGSGTNYDACHREGF

Signal sequence:

amino acids 1-18

FIGURE 156

AAAAAAGCTCACTAAAGTTTCTATTAGAGCGAATACGGTAGATTTCATCCCCCTTTTGAAGAACAGTACTGTGGA
GCTATTTAAGAGATAAAAACGAAATATCCTTTCTGGGAGTTCAAGATTGTGCAGTAAATTGGTTAGGACTCTGAGC
GCCGCTGTTTACCAATCGGGGAGAGAAAAGCGGAGATCCTGCTCGCCTTGACGCGCCTGAAGCACAAAGCAGAT
AGCTAGGAATGAACCATCCCTGGGAGTATGTGGAACAACGGAGGAGCTCTGACTTCCCAACTGTCCCATTCTAT
GGGCGAAGGAACTGCTCCTGACTTCAGTGGTTAAGGGCAGAATTGAAAATAATTCTGGAGGAAGATAAGAAATGAT
TCCTGCGCGACTGCACCGGACTACAAAGGGCTTGTCTGCTGGGAATCCTCCTGGGGACTCTGTGGGAGACCGG
ATGCACCCAGATACGCTATTCACTTCCGGAAGAGCTGGAGAAAGGCTCTAGGGTGGGCGACATCTCCAGGGACCT
GGGGCTGGAGCCCCGGGAGCTCGCGGAGCGCGGAGTCCGCATCATCCCCAGAGGTAGGACGCAGCTTTTCGCCCT
GAATCCGCGCAGCGGCAGCTTGGTCACGGCGGGCAGGATAGACCGGGAGGAGCTCTGTATGGGGGCCATCAAGTG
TCAATTAAATCTAGACATTCTGATGGAGGATAAAGTGAAAATATATGGAGTAGAAGTAGAAGTAAGGGACATTAA
CGACAATGCGCCTTACTTTCTGTGAAAGTGAATTAGAAAATAAAAATTAGTGAAAATGCAGCCACTGAGATGCGGTT
CCCTCTACCCACGCTCGGATCCGATATCGGGAAGAACTCTCTGCAGAGCTACGAGCTCAGCCCCGAACACTCA
CTTCTCCCTCATCGTGCAAAATGGAGCCGACGGTAGTAAGTACCCGAATTGGTGCTGAAACGCGCCCTGGACCG
CGAAGAAAAGGCTGCTCACCACCTGGTCTTACGGCCTCCGACGGGGCGACCCGGTGCGCACAGGCACCGCGG
CATCCGCGTGATGGTTCTGGATGCGAACGACAACGCACCCAGCGTTTGCTCAGCCGAGTACCGCGCAGCGTTCC
GGAGAATCTGGCCTTGGGCACGCAGCTGCTTGTAGTCAACGCTACCGACCTGACGAAGGAGTCAATGCGGAAGT
GAGGTATTCTTCCGGTATGTGGACGACAAGGCGGCCCAAGTTTCAAACCTAGATTGTAATTACGGGACAATATC
AACAATAGGGGAGTTGGACCACGAGGAGTCAGGATTCTACCAGATGGAAGTGCAAGCAATGGATAATGCAGGATA
TTCTGCGCGAGCCAAAGTCTTGATCACTGTTCTGGACGTGAACGACAATGCCCCAGAAGTGGTCTCACCTCTCT
CGCCAGCTCGGTTCCCGAAAACCTCTCCAGAGGGACATTAATTGCCCTTTTAAATGTAAATGACCAAGATTCTGA
GGAAAACGGACAGGTGATCTGTTTCAATCCAGGAAATCTGCCCTTTAAATTAGAAAATCTTACGGAAATTACTA
TAGTTTATGCACAGACATAGTCTTGGATAGGGAAACAGGTTCTAGCTACAACATCACAGTGACCGCCACTGACCG
GGGAACCCCGCCCCATCCACGGAACCTCATATCTGCTGAACGTGGCAGACACCAACGACAACCCGCGCGTCTT
CCCTCAGGCCTCCTATTCCGCTTATATCCCAGAGAACAATCCCAGAGGAGTTTCCCTCGTCTCTGTGACCGCCCA
CGACCCGACTGTGAAGAGAACGCCAGATCACTTATTCCCTGGCTGAGAACACCATCCAAGGGGCAAGCCTATC
GTCCTACGTGTCCATCAACTCCGACACTGGGGTACTGTATGCGCTGAGCTCCTTCGACTACGAGCAGTTCCGAGA
CTTGCAAGTGAAAGTGATGGCGCGGACAAACGGGCACCCGCCCTCAGCAGCAACGTGTCTGTTGAGCCTGTTCTG
GCTGGACCAGAACGACAATGCGCCCGAGATCCTGTACCCCGCCCTCCCCACGGACGGTTCCTGCGTGGAGCT
GGCTCCCCGCTCCGACAGAGCCCGGCTACCTGGTGACCAAGGTGGTGGCGGTGGACAGAGACTCCGGCCAGAAGC
CTGGCTGTCTACCGTCTGCTCAAGGCCAGCGAGCCGGACTCTTCTCGGTGGGTCTGCACACGGGCGAGGTGCG
CACGGCGCGAGCCCTGCTGGACAGAGACGCGCTCAAGCAGAGCCTCGTAGTGGCCGTCCAGGACCAGGCCAGCC
CCCTCTCTCCGCCACTGTACGCTCACCCTGGCCGTGGCCGACAGCATCCCCAAGTCTTGGCGGACCTCGGCAG
CCTCGAGTCTCCAGCTAACTCTGAAACCTCAGACCTCACTCTGTACCTGGTGGTAGCGGTGGCCGCGGTCTCCTG
CGTCTTCTGGCCTTCTGTCATCTTGTCTGCTGGCGCTCAGGCTGCGGCGCTGGCACAAGTCACGCCTGCTGCAGGC
TTCAGGAGGCGGCTTGACAGGAGCGCCGGCGTGCACCTTTGTGGGCGTGGACGGGGTGCAGGCTTTTCTGCAGAC
CTATTCCCACGAGGTTTCCCTCACCACGGACTCGCGGAAGAGTCACCTGATCTTCCCCCAGCCCAACTATGCAGA
CATGCTCGTCAGCCAGGAGAGCTTTGAAAAAAGCGAGCCCTTTTGCTGTGAGGTGATTGGGTATTTTCTAAAGA
CAGTCATGGGTAAATTGAGGTGAGTTTATATCAAATCTTCTTTCTTTTCTTTTAAATTGCTCTGTCTCCCAAGC
TGGAGTGACGCGGTACGATCATAGCTCACTGCGGCTCAAACCTCCTAGGCTCAAGCAATTATCCCACCTTTGCCT
CCGGTGTAACAGGGACTACAGGTGCAAGCCACCTACTGTCTGCCTATCTATCTATCTATCTATCTATCTATCTAT
CTATCTATCTATCTATCTATTACTTTCTTGTACAGACGGGAGTCTCACGCCTGTAATCCCAGTACTTTGGGAGGC
CGAGGCGGGTGGATCACCTGAGGTTGGGAGTTTGAGACCAGCTGACCAACATGGAGAAACCCGCTCTATACTAA
AAAAATACAAATTAGCCGGGCGTGGTGGTGCAATGCTGTAAATCCCAGTACTTGGGAGGCTGAGTCAGGAGAAT
TGCTTTAACCTGGGAGGTGGAGGTGCAATGAGCTGAGATTGTGCCATTGCACTCCAGCCTGGGCAACAAGAGTG
AAACTCTATCTCA

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306

><subunit 1 of 1, 916 aa, 1 stop

><MW: 100204, pI: 4.92, NX(S/T): 4

MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVPEELEKGSRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKISENAATEMRFPPLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELVLKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAPQPEYR
ASVPENLALGTQLLVVNATDPDEGVNAEVRYSFRYVDDKAAQVFKLDCNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLNVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYYSLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLNVADTNDNPPVFPQASYSAYIPENNPRGVSLSVSTAHDPCCEENAQITYSLAENTI
QGASLSSYVSINSDTGVLYALSSFDYEQFRDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDN
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRDSGQNAWLSYRLLKASEPGLFSVG
LHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTTLTVAVADSIPQVLADLGSLESPA
NSETSDLTLYLVVAVAAVSCVFLAFVILLALRLRRWHKSRLQASGGGLTGAPASHFVGVD
GVQAFLOQTYSHEVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGL
IEVSLYQIFFLFFFNCVSQAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRDYZRCKPPTVCLS
IYLSIYLSIYLSIYLLSCTDGS LTPVIPVLWEAEAGGSPEVGSLRPA

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAG
GCTCCAGAATCGTGTAACAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCCGCCCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCTTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGGTCAAGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGG
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAAGGTGGCAACTGGGTCTTT
ACAGCTGCCCACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA
TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC
CTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTG
CACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACT
GTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGG
CCCCCTGGTGTGTGATGGTGCACCTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACCTCT
CTGGTTC

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336

<subunit 1 of 1, 260 aa, 1 stop

<MW: 28048, pI: 7.87, NX(S/T): 1

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVL
VGGNWLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSSGGLVCDGALQGITSWGS DPCGRSDKPGVYTNICRY
LDWIKKIIGSKG

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

FIGURE 160

GGCGCCGGTGCACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCTGCGCGCCCCGGCCCCG
CGCGCCCGCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCCGGGCCCGCGCCCCG
GCCCCGCGCCAGGTGAGCGCTCCGCCCGCCGCGAGGCCCGCCCCGGCCCCCGCCCCGCCCCG
CCCCGGCCGGCGGGGGAACCGGGCGGATTCTCTGCGCGTCAAACCACCTGATCCCATAAAAC
ATTCACTCTCCCGGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCGCCGCGCCCTCG
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCAGCCAGAGCCGGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCGCGGCCCGGGCCGGGGCCGGGCCGTAGCGGCGGCGCCTGGA
TGCGGACCCCGCCGCGGGGAGACGGGCGCCCCGCCCGGAAACGACTTTCAGTCCCCGACGCGC
CCCCGCCAACCCCTACGATGAAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTG
CTGTGGCTGCAGGCCTGGCAGGTGGCAGCCCCATGCCAGGTGCCTGCGTATGCTACAATGA
GCCCCAAGGTGACGACAAGCTGCCCCCAGCAGGGCCTGCAGGCTGTGCCCGTGGGCATCCCTG
CTGCCAGCCAGCGCATCTTCTGACGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC
CGTGCTGCGCAACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCGAATTGATGC
GGCTGCCTTCACTGGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCCTGCCACATTCACGGCCTGGGCCCGCTACACACGCTGCACCTGGACCGC
TGCGGCCTGCAGGAGCTGGGCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACGCGCTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCA
CACACCTCTTCTGACGGCAACCGCATCTCCAGCGTGCCCGAGCGCGCCTTCCGTGGGCTG
CACAGCCTCGACCGTCTCCTACTGCACCAGAACCGCGTGGCCCATGTGCACCCGCATGCCTT
CCGTGACCTTGGCCGCCTCATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCTGCCCA
CTGAGGCCCTGGCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTG
TGTGACTGCCGGGCACGCCCACTCTGGGCCTGGCTGCAGAAGTTCGCGGGCTCCTCCTCCGA
GGTGCCCTGCAGCCTCCCGCAACGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATG
ACCTGCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACC
GATGAGGAGCCGCTGGGGCTTCCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGT
ACTGGAGCCTGGAAGACCAGCTTTCGGCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCCGGTG
ACAGCCCCGCGGGCAACGGCTCTGGCCACGGCACATCAATGACTCACCTTTGGGACTCTG
CCTGGCTCTGCTGAGCCCCCGCTCACTGCAGTGCGGCCCGAGGGCTCCGAGCCACCAGGGTT
CCCCACCTCGGGCCCTCGCCGGAGGCCAGGCTGTTACGCAAGAACCGCACCCGCAGCCACT
GCCGTCTGGGCCAGGCAGGCAGCGGGGTGGCGGGACTGGTGACTCAGAAGGCTCAGGTGCC
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGGCCTGGCGCTGGTGCTGTGGACAGTGCT
TGGGCCCTGCTGAACCCCCAGCGGACACAAGAGCGTGCTCAGCAGCCAGGTGTGTGTACATAC
GGGGTCTCTCTCCACGCCGCAAGCCAGCCGGGCGGCCGACCCGTGGGGCAGGCCAGGCCAG
GTCTCTCCCTGATGGACGCCTGCCGCCGCCACCCCCATCTCCACCCCATCATGTTTACAGGG
TTCGGCGGCAGCGTTTGTTCAGAACGCCGCTCCACCCAGATCGCGGTATATAGAGATAT
GCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAAGAGCTCTTTTCTTAAA
AAAA

FIGURE 161

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184

><subunit 1 of 1, 473 aa, 1 stop

><MW: 50708, pI: 9.28, NX(S/T): 6

MKRASAGGSRL LAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQOGLQAVPVGIPAASQRI
FLHG NRISHVPAASFRACRNLTILWLHSNVLARIDAAFTGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLLHQNRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRRLNDNPWVCDRCRARPLWAWLQKFRGSSSEVP CSLPQRLAGRDLKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDAAADKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSEGS GALPSLTCSLTPLGLALVLWTVLGPC

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

FIGURE 162

GGAAGTCCACGGGGAGCTTGGATGCCAAAGGGAGGACGGCTGGGTCTCTGGAGAGGACTAC
TCACTGGCATATTTCTGAGGTATCTGTAGAATAAACCACAGCCTCAGATACTGGGGACTTTAC
AGTCCACACAGAACCGTCTCTCCAGGAAGCTGAATCCAGCAAGAACAATGGAGGCCAGCGGGA
AGCTCATTTGCAGACAAAGGCAAGTCCTTTTTTCTCTTTTGGGCTTATCTCTGGCG
GGCGCGGCGGAACCTAGAAAGCTATTCTGTGGTGGAGGAAACTGAGGGCAGCTCCTTTGTAC
CAATTTAGCAAAGGACCTGGGTCTGGAGCAGAGGGAATTCTCCAGGCGGGGGGTTAGGGTTG
TTTCCAGAGGGAAACAACTACATTTGCAGCTCAATCAGGAGACCGCGGATTTGTTGCTAAAT
GAGAAATTGGACCGTGAGGATCTGTGCGGTACACAGAGCCCTGTGTGCTACGTTTCCAAGT
GTTGCTAGAGAGTCCCTTCGAGTTTTTTCAAGCTGAGCTGCAAGTAATAGACATAAACGACC
ACTCTCCAGTATTTCTGGACAAACAAATGTTGGTGAAAGTATCAGAGAGCAGTCTCTCTGGG
ACTACGTTTCTCTGAAGAATGCCGAAGACTTAGATGTAGGCCAAAACAATATTGAGAACTA
TATAATCAGCCCCAACTCCTATTTTTCGGGTCTCACCCGCAAACGCAGTGATGGCAGGAAAT
ACCCAGAGCTGGTGCTGGACAAAGCGCTGGACCGAGAGGAAGAAGCTGAGCTCAGGTTAAACA
CTCACAGCACTGGATGGTGGCTCTCCGCCCAGATCTGGCACTGCTCAGGTCTACATCGAAGT
CCTGGATGTCAACGATAATGCCCTGAATTTGAGCAGCCTTTCTATAGAGTGCAGATCTCTG
AGGACAGTCCGGTAGGCTTCCTGGTTGTGAAGGTCTCTGCCACGGATGTAGACACAGGAGTC
AACGGAGAGATTTCTTATTTCACTTTTTCCAAGCTTCAGAAGAGATTGGCAAACCTTTAAGAT
CAATCCCTTGACAGGAGAAATTGAACTAAAAAACAACCTCGATTTCGAAAACTTCAGTCCT
ATGAAGTCAATATTGAGGCAAGAGATGCTGGAACCTTTTTCTGGAAAATGCACCGTTCTGATT
CAAGTGATAGATGTGAACGACCATGCCCCAGAAGTTACCATGTCTGCATTTACCAGCCCAAT
ACCTGAGAACGCGCCTGAAACTGTGGTTGCACTTTTTCAGTGTTTCAGATCTTGATTTCAGGAG
AAAATGGGAAAATTAGTTGCTCCATTTCAGGAGGATCTACCTTTCTCTGAAATCCGCGGAA
AACTTTTACACCCTACTAACGGAGAGACCACTAGACAGAGAAAGCAGAGCGGAATACAACAT
CACTATCACTGTCACTGACTTGGGGACCCCTATGCTGATAACACAGCTCAATATGACCGTGC
TGATCGCCGATGTCAATGACAACGCTCCCGCCTTCACCCAAACCTCCTACACCCTGTTTCGTC
CGCGAGAACAACAGCCCCGCCCTGCACATCCGCAGCGTCAGCGCTACAGACAGAGACTCAGG
CACCAACGCCCAGGTACCTACTCGCTGCTGCCGCCCCAGGACCCGCACCTGCCCTCACAT
CCCTGGTCTCCATCAACGCGGACAACGGCCACCTGTTCCGCCCTCAGGTCTCTGGACTACGAG
GCCCTGCAGGGGTTCCAGTTCCGCGTGGGCGCTTCAGACCACGGCTCCCCGGCGCTGAGCAG
CGAGGCGCTGGTGCGCGTGGTGGTGCTGGACGCCAACGACAACCTCGCCCTTCGTGCTGTACC
CGCTGCAGAACGGCTCCGCGCCCTGCACCGAGCTGGTGCCCCGGGCGGCCGAGCCGGGCTAC
CTGGTGACCAAGGTGGTGGCGGTGGACGCGCACTCGGGCCAGAACGCCTGGCTGTCTGTACCA
GCTGCTCAAGGCCACGGAGCTCGGTCTGTTCGGCGTGTGGGCGCACAATGGCGAGGTGCGCA
CCGCCAGGCTGCTGAGCGAGCGCGACGCGGCCAAGCACAGGCTGGTGGTGCTGGTCAAGGAC
AATGGCGAGCCTCCGCGCTCGGCCACCGCCACGCTGCACGTGCTCCTGGTGGACGGCTTCTC
CCAGCCCTACCTGCCTCTCCCGGAGGCGGCCCGACCCAGGCCAGGCCGACTTGCTCACCG
TCTACCTGGTGGTGGCGTTGGCCTCGGTGTCTTCGCTCTTCCTCTTTTCGGTGCTCCTGTTT
GTGGCGGTGCGGCTGTGTAGGAGGAGCAGGGCGGCCCTCGGTGGGTGCTGCTTGGTGCCCGA
GGGCCCCCTTCCAGGGCATCTTGTGGACATGAGCGGCACCAGGACCCTATCCCAGAGCTACC
AGTATGAGGTGTGTCTGGCAGGAGGCTCAGGGACCAATGAGTTCAAGTTCCTGAAGCCGATT
ATCCCCAACTTCCCTCCCCAGTGCCCTGGGAAAGAAATACAAGGAAATTCTACCTTCCCCAA
TAACTTTGGGTTCAATATTCAGTGACCATAGTTGACTTTTACATTCCATAGGTATTTTATTT
TGTGGCATTTCATGCCAATGTTTATTTCCCCCAATTTGTGTGTATGTAATATTGTACGGAT
TTACTCTTGATTTTTCTCATGTTCTTTCTCCCTTGTTTTAAAGTGAACATTTACCTTTATT
CCTGGTTCTT

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314

<subunit 1 of 1, 798 aa, 1 stop

<MW: 87552, pI: 4.84, NX(S/T): 5

MEASGKLICRQRQVLFSFLLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSARGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRQVLLES PF EFFQ AELQV
IDINDHSPVFLDKQMLVKVSESSPPGTTTFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLDVNDNAPEFEQPFY
RVQISEDSPVGFLVVKVSATDVD TG VNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAP E V T M S A F T S P I P E N A P E T V V A L F S V S
DLDSGENGKISCSIQEDLPFLLKSAENFYTLLTERPLDRESRAEYNITITVTDLGTPMLITQ
LNMTVLIADVNDNAPAFQTQTSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELG L F G V W A H
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAA
ADLLTVYLVVALASVSSLFLFSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMSGTRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFNIQ

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

FIGURE 164

ACCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCGCGTAGCCGTGC
GCCGATTGCCTCTCGGCCTGGGCAATGGTCCCGGCTGCCGGTCGACGACCGCCCCGCGTCAT
GCGGCTCCTCGGCTGGTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCCGTCCGCGGCGTG
AGGTTGCAGAGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCCTCTCCAGGTG
GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGCCAGGACAGGGCAGC
AGAAGAGGCCAATGCGGTGCTGGGGCTGGACACCCAAGGCGATCACATGGTGATGCTGTCTG
TGATTCTGGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGGCGTCACCTGTGGTGCT
GGAGGAGCGGAGGACTCAAGGTGCAACGTCCGAGAGAGCCTTTTCTCTCTGGATGGCGCTGG
AGCACACTTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG
CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACTGTGAG
GAGAGAAACATTACAGGATTAGAAAATTTCACTCTGAAAATTTTAAATATGTACAGGACCT
TATGGATTTTCTGAACCCAAACGGTAGTGACTGTACTCTAGTCCTGTTTTACACCCCGTG
GCCGCTTTTCTGCCAGTTTGGCCCCCTCACTTTAACTCTCTGCCCCGGGCATTTCCAGCTCTT
CACTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTTCTACCAGGTTTGGCACCGTAGC
TGTTCCATAATATTTTATTATTTCAAGGAGCTAAACCAATGGCCAGATTTAATCATACAGATC
GAACACTGGAAACACTGAAAATCTTCATTTTTAATCAGACAGGTATAGAAGCCAAGAAGAAT
GTGGTGGTAACTCAAGCCGACCAAAATAGGCCCTCTCCAGCACTTTGATAAAAAGTGTGGA
CTGGTTGCTTGTATTTTCCTTATTCTTTTAAATTAGTTTTATTATGTATGCTACCATTTCGAA
CTGAGAGTATTCGGTGGCTAATTCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
GAAAGAAGTTGGAAAGAGGAACCTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTTTCATA
CATTTTCTCCAGTGACGTGTTGACTTGAAACTTCAGGCAGATTAAAAGAATCATTTGTTGAA
CAACTGAATGTATAAAAAAATTATAAACTGGTGTTTTAACTAGTATTGCAATAAGCAAATGC
AAAAATATTCAATAG

FIGURE 165

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333

><subunit 1 of 1, 360 aa, 1 stop

><MW: 39885, pI: 4.79, NX(S/T): 7

MVPAAGRRPVRMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEE
ELLHDPMGQDRAAEANAVLGLDTQGDHVMVLSVIPGEAEDKVSSEPSGVTCGAGGAEDSRC
NVRESLFLSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLE
NFTLKILNMSQDLMDFLNPNNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPALHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMAFNFHTDRTLETLKIFIFNQGTGIEAKKNVVVTQADQ
IGPLPSTLIKSDWLLVFSLFFLISFIMYATIRTESIRWLIPGQEQEHVE

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTCGCGCGCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTGCTCTTCCTCGCCTCGCACTGCTGCCTGGGCTCGGCGCGGGGCTCTTCCTC
TTTGGCCAGCCCCGACTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAGAACATGCGGGCTGCCCAACCTGCTGGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCG
GACACCAAGAAGTTCTGTGCTCGCTCTTCGCCCCCGTCTGCCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCTGCGTGCAGGTGAAGGACCGCTGCGCCCCGGTCATGTCCG
CCTTCGGCTTCCCCCTGGCCCGACATGCTTGAGTGCAGCCGTTTCCCCCAGGACAACGACCTT
TGCATCCCCCTCGCTAGCAGCGACCACCTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATG
TGAAGCCTGCAAAAATAAAAATGATGATGACAACGACATAATGGAAACGCTTTGTAAAAATG
ATTTTGCACTGAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAATCATC
CTGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGCACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCATGGGACAGAAACAGGGTGGGGAGCTGGTGATCACCTCGGTGAAGCGG
TGGCAGAAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGCTTA
GTCCCGGCATCCTGATGGCTCCGACAGGCCTGCTCCAGAGCACGGCTGACCATTTCTGCTCC
GGGATCTCAGCTCCCGTTCCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGGCAGCT
TCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTCTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTTACCTAAAGGAAAAGCCACCCGAATCTTGTAGAAATATTCAAACATAATA
AAATCATGAATATTTTAA

FIGURE 167

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920

><subunit 1 of 1, 295 aa, 1 stop

><MW: 33518, pI: 7.74, NX(S/T): 0

MLQGPGLLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHS LCVQVKDR
CAPVMSAFGFPPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDNDNDIM
ETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLVGVSEKDLKKS LKSLQCTCE
EMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISR SIRKLQC

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus

amino acids 6-153

167-168-169-170-171-172-173-174-175-176-177-178-179-180-181-182-183-184-185-186-187-188-189-190-191-192-193-194-195-196-197-198-199-200-201-202-203-204-205-206-207-208-209-210-211-212-213-214-215-216-217-218-219-220-221-222-223-224-225-226-227-228-229-230-231-232-233-234-235-236-237-238-239-240-241-242-243-244-245-246-247-248-249-250-251-252-253-254-255-256-257-258-259-260-261-262-263-264-265-266-267-268-269-270-271-272-273-274-275-276-277-278-279-280-281-282-283-284-285-286-287-288-289-290-291-292-293-294-295-296-297-298-299-300-301-302-303-304-305-306-307-308-309-310-311-312-313-314-315-316-317-318-319-320-321-322-323-324-325-326-327-328-329-330-331-332-333-334-335-336-337-338-339-340-341-342-343-344-345-346-347-348-349-350-351-352-353-354-355-356-357-358-359-360-361-362-363-364-365-366-367-368-369-370-371-372-373-374-375-376-377-378-379-380-381-382-383-384-385-386-387-388-389-390-391-392-393-394-395-396-397-398-399-400-401-402-403-404-405-406-407-408-409-410-411-412-413-414-415-416-417-418-419-420-421-422-423-424-425-426-427-428-429-430-431-432-433-434-435-436-437-438-439-440-441-442-443-444-445-446-447-448-449-450-451-452-453-454-455-456-457-458-459-460-461-462-463-464-465-466-467-468-469-470-471-472-473-474-475-476-477-478-479-480-481-482-483-484-485-486-487-488-489-490-491-492-493-494-495-496-497-498-499-500-501-502-503-504-505-506-507-508-509-510-511-512-513-514-515-516-517-518-519-520-521-522-523-524-525-526-527-528-529-530-531-532-533-534-535-536-537-538-539-540-541-542-543-544-545-546-547-548-549-550-551-552-553-554-555-556-557-558-559-560-561-562-563-564-565-566-567-568-569-570-571-572-573-574-575-576-577-578-579-580-581-582-583-584-585-586-587-588-589-590-591-592-593-594-595-596-597-598-599-600-601-602-603-604-605-606-607-608-609-610-611-612-613-614-615-616-617-618-619-620-621-622-623-624-625-626-627-628-629-630-631-632-633-634-635-636-637-638-639-640-641-642-643-644-645-646-647-648-649-650-651-652-653-654-655-656-657-658-659-660-661-662-663-664-665-666-667-668-669-670-671-672-673-674-675-676-677-678-679-680-681-682-683-684-685-686-687-688-689-690-691-692-693-694-695-696-697-698-699-700-701-702-703-704-705-706-707-708-709-710-711-712-713-714-715-716-717-718-719-720-721-722-723-724-725-726-727-728-729-730-731-732-733-734-735-736-737-738-739-740-741-742-743-744-745-746-747-748-749-750-751-752-753-754-755-756-757-758-759-760-761-762-763-764-765-766-767-768-769-770-771-772-773-774-775-776-777-778-779-780-781-782-783-784-785-786-787-788-789-790-791-792-793-794-795-796-797-798-799-800-801-802-803-804-805-806-807-808-809-810-811-812-813-814-815-816-817-818-819-820-821-822-823-824-825-826-827-828-829-830-831-832-833-834-835-836-837-838-839-840-841-842-843-844-845-846-847-848-849-850-851-852-853-854-855-856-857-858-859-860-861-862-863-864-865-866-867-868-869-870-871-872-873-874-875-876-877-878-879-880-881-882-883-884-885-886-887-888-889-890-891-892-893-894-895-896-897-898-899-900-901-902-903-904-905-906-907-908-909-910-911-912-913-914-915-916-917-918-919-920-921-922-923-924-925-926-927-928-929-930-931-932-933-934-935-936-937-938-939-940-941-942-943-944-945-946-947-948-949-950-951-952-953-954-955-956-957-958-959-960-961-962-963-964-965-966-967-968-969-970-971-972-973-974-975-976-977-978-979-980-981-982-983-984-985-986-987-988-989-990-991-992-993-994-995-996-997-998-999-1000

FIGURE 168

GTGGAGGCCGCCGACGATGGCGGGGCCGACGGAGGCCGAGACGGGGTTGGCCGAGCCCCGGG
CCCTGTGCGCGCAGCGGGGCCACCGCACCTACGCGCGCCGCTGGGTGTTCTTGCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGCTGACGTGAT
TGCTGAGGACTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCCATTTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTGCGGCTCCGTGCGGCG
ACCATCCTGGGTGCGTGGCTGAACTTTGCCGGGAGTGTGCTACGCATGGTGCCCTGCATGGT
TGTTGGGACCCAAAACCCATTTGCCTTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTTGCCC
AGAGCCTGGTCATCTTCTCTCCAGCCAAGCTGGCTGCCTTGTGGTTCCAGAGCACCAGCGA
GCCACGGCCAACATGCTCGCCACCATGTGGAACCTCTGGGCGTCTTGTGGCCAATGTGCT
GTCCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTCGTCTGCCTGCTGTCCACCATCTGCCTGTGGGAGAGTGTGCCCCCACC
CCCTCTGCCGGGGCTGCCAGCTCCACCTCAGAGAAGTTCTGGATGGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTCATCCTGGCTGTGTGCTTGGGGGAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGGTTTTCC
GGCCTCTGTGGCGCTCTCTTCATCACGTTTGGGATCCTGGGGGCACTGGCTCTCGGCCCTA
TGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTCTGG
CCTGCGTGCCCTTTGCCCTGGTGTCCCAGCTGCAGGGACAGACCCTTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGGCTGTTTGGCTTCTCGGTGGGCCCCGTGGCCATGGAGTTGGCGGTGGA
GTGTTCTCTTCCCCGTGGGGGAGGGGGCTGCCACAGGCATGATCTTTGTGCTGGGGCAGGCCG
AGGGAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGTCTTG
TCCACCTGCCAGCAGGGGGAGGATCCACTTGACTGGACAGTGTCTCTGCTGCTGATGGCCGG
CCTGTGCACCTTCTTCAGCTGCATCCTGGCGGTCTTCTTCCACACCCCATACCGGCGCCTGC
AGGCCGAGTCTGGGGAGCCCCCTCCACCCGTAACGCCGTGGGCGGCGCAGACTCAGGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGGTCTGGGGCCAGCACGGCGACTCCGGA
GTGCACGGCGAGGGGGGCTCGCTAGAGGACCCAGAGGGCCCGGAGCCCCACCCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCAGCAGCCACCGACGCGCCCTCCCGCCCCGGC
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTTATTGACCCGGCTGGGTCTCACTCCTCCTT
CTCCTCCCCGTGGGTGATCACGTAGCTGAGCGCCTTGTAGTCCAGGTTGCCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGGTCCACCTGCGGGCGGGGGCGAAAGGGCTCCTTGCGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988

><subunit 1 of 1, 560 aa, 1 stop

><MW: 58427, pI: 6.86, NX(S/T): 2

MAGPTEAETGLAEPRALCAQRGHRITYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAAIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVGTON
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQ RATANMLATMSNPLGVLVANVLSPVLV
KKGEDIPLMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNKA
YVILAVCLGGMIGISASFSALEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCFLSLACVPFALVSQLOGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GEGAATGMIFVLGQAEGLIMLAMTALTVRREPSLSTCQQGEDPLDWTVSLLL MAGLCTFF
SCILAVFFHTPYRRLQAESGEPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG
ASLEDPRGPGSPHPACHRATPRAQGPAA TDAPSRPGRLAGRVQASRFIDPAGSHSSFSSPWVIT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

FIGURE 170

GTCCACATCCTGCTCAACTGGGTGAGGTCCCTCTTAGACCAGCTCTTGTCATCATTTGCTGAAGTGGACCAAC
TAGTTCCCCAGTAGGGGGTCTCCCTGGCAATTCTTGATCGGCGTTTGGACATCTCAGATCGCTTCCAATGAAGA
TGGCCTTGCCCTTGGGGTCCCTGCTTGTTCATAATCATCTAACTATGGGACAAGGTTGTGCCGGCAGCTCTGGGGG
AAGGAGCACGGGGCTGATCAAGCCATCCAGGAAACACTGGAGGACTTGTCCAGCCTTGAAAGAACTCTAGTGGTT
TCTGAATCTAGCCCACTTGGCGGTAAGCATGATGCAACTTCTGCAACTTCTGCTGGGGCTTTTGGGGCCAGGTGG
CTACTTATTTCTTTTAGGGGATGTGTCAGGAGGTGACCACTCTCACGGTGAAATACCAAGTGTGTCAGAGGAAGTGCC
ATCTGGTACAGTGATCGGGAAGCTGTCCCAGGAACTGGGCGGGAGGAGAGGCGGAGGCAAGCTGGGGCCGCTT
CCAGGTGTTGCAGCTGCCCTCAGGCGCTCCCCATTACGGTGGACTCTGAGGAAGGCTTGCTCAGCACAGGCAGGCG
GCTGGATCGAGAGCAGCTGTGCCGACAGTGGGATCCCTGCTGGTTTCTTTGATGTGCTTGCCACAGGGGATTT
GGCTCTGATCCATGTGAGATCCAAGTGCTGGACATCAATGACCACCAGCCACGGTTTCCCAAAGGCGAGCAGGA
GCTGGAAATCTCTGAGAGCGCTCTCTGCGAACCCGGATCCCCCTGGACAGAGCTCTTGACCCAGACACAGGCCC
TAACACCCTGCACACCTACACTCTGTCTCCCAGTGAGCACTTTGCCTTGATGTGCTTGTGGGCCCTGATGAGAC
CAAACATGCAGAACTCATAGTGGTGAAGGAGCTGGACAGGGAAATCCATTCATTTTTTGATCTGGTGTAACTGC
CTATGCAATGGGAACCCCCCAAGTCAGGTACCACTTGGTCAAGGTCAACGTCTTGGACTCCAATGACAATAG
CCTTGCCTTTGCTGAGAGTTCACTGGCACTGGAATCCAAGAAGATGCTGCACCTGGTACGCTTCTCATAAAAC
GACCGCCACAGCCCTGACCAAGGCCCAATGGGGAGGTGGAGTTCTTCTCAGTAAGCACATGCCTCCAGAGGT
GCTGGACACCTTCAGTATTGATGCCAAGACAGGCCAGGTCACTTCTGCTCGACCTCTAGACTATGAAAAGAACCC
TGCCCTACGAGGTGGATGTTTCAAGCAAGGACCTGGGTCCCAATCCTATCCAGCCCATTCGAAAAGTTCTCATCAA
GGTTCTGGATGTCAATGACAACATCCCAAGCATCCACGTCAATGGGCTCCAGCCATCACTGGTGTGAGAAGC
TCTTCCCAAGGACAGTTTATTGCTCTTGTGTCATGGCAGATGACTTGGATTGAGGACACAATGGTTTGGTCCACTG
CTGGCTGAGCCAAGAGCTGGGCCACTTCAAGCTGAAAAGAATAATGGCAACACATACATGTTGCTAACCAATGC
CACACTGGACAGAGAGCAGTGGGCCAAATATACCTCACTCTGTTAGCCCAAGACCAAGGACTCCAGCCCTTATC
AGCCAAGAAACAGCTCAGCACTCAGATCAGTGACATCAACGACAATGCACCTGTGTTTGAGAAAAGCAGGTATGA
AGTCTCCACGCGGGAAAACAACCTTACCCTCTCTTACCTCATTACCATCAAGGCTCATGATGCAGACTTGGGCAT
TAATGGAAAAGTCTCATACCGCATCCAGGACTCCCCAGTTGCTCACTTAGTAGCTATTGACTCCAACACAGGAGA
GGTCACTGCTCAGAGGTCACTGAACATGAAGAGATGGCCGGCTTTGAGTTCAGGTGATCGCAGAGGACAGCGG
GCAACCCATGCTTGCAATCCAGTGTCTCTGTGTGGGTCAAGCTCTTGATGCCAATGATAATGCCCCAGAGGTGGT
CCAGCCTGTGCTCAGCGATGGAAAAGCCAGCCTCTCCGTGCTTGTGAATGCCTCCACAGGCCACCTGCTGGTGCC
CATCGAGACTCCCAATGGCTTGGGCCCAGCGGGCACTGACACACCTCCACTGGCCACTCACAGCTCCCGGCCATT
CCTTTTGACAACCAATTGTGGCAAGAGATGCAGACTCGGGGGCAAATGGAGAGCCCCTCTACAGCATCCGCAATGG
AAATGAAGCCCACCTCTTCATCTCAACCCCTCATACGGGGCAGCTGTTTGTCAATGTACCAATGCCAGCAGCCT
CATTTGGGAGTGAGTGGGAGCTGGAGATAGTAGAGGACCAGGGAAGCCCCCTTACAGACCCGAGCCCTGTT
GAGGTCATGTTTGTACCCAGTGTGGACCCTTGAGGGAAGTCAAGCCGCAAGCCTGGGGCCCTTGAGCATGTCGAT
GCTGACGGTGATCTGCCTGGCTGTACTGTGGGCATCTTGGGTTGATCCTGGCTTTGTTTATGTCCATCTGCCG
GACAGAAAAGAAGGACAACAGGGCCTACAACCTGTGCGGAGGCGGAGTCCACCTACCGCCAGCAGCCCAAGAGGCC
CCAGAAACACATTGAGAAGGCAGACATCCACCTCGTGCTGTGCTCAGGGGTGAGGAGGTGAGCCTTGTGAAGT
CGGGCAGTCCCAAAAGATGTGGACAAGGAGGCGATGATGGAAGCAGGCTGGGACCCCTGCCTGCAGGCCCCCTT
CCACCTCACCCCGACCTGTACAGGACGCTGCGTAATCAAGGCAACCAGGGAGCACCGGCGGAGAGCCGAGAGGT
GCTGCAAGACACGGTCAACCTCCTTTTCAACCATCCCAAGGAGGGAATGCCTCCCGGAGAACCTGAACTTCC
CGAGCCCCAGCCTGCCACAGGCCAGCCACGTTCCAGGCCTCTGAAGGTTGCAGGCAGCCCCACAGGGAGGCTGGC
TGGAGACCAGGGCAGTGAGGAAGCCCCACAGAGGCCACAGCCTCCTCTGCAACCTGAGACGGCAGCGACATCT
CAATGGCAAAGTGTCCCTGAGAAAGAATCAGGGCCCCGTGAGATCCTGCGGAGCCTGGTCCGGCTGTCTGTGGC
TGCCTTCGCGGAGCGGAACCCGTGGAGGAGCTCACTGTGGATTCTCCTCCTGTTTCAAGAAATCTCCAGCTGCT
GTCCTTGTGTCATCAGGGCCAATTCCAGGCCAAACCAACACCGAGGAAATAAGTACTTGGCCAAGCCAGGAGG
CAGCAGGAGTGCAATCCAGACACAGATGGCCCAAGTGCAAGGGCTGGAGGCCAGACAGACCCAGAACAGGAGGA
AGGGCCTTTGGATCTGAAGAGGACCTCTCTGTGAAGCAACTGCTAGAAGAAGAGCTGTCAAGTCTGCTGGACCC
CAGCACAGGTCTGGCCCTGGACCGGCTGAGCGCCCTGACCCGGCCTGGATGGCGAGACTCTCTTTGCCCTCAC
CACCACCTACCGTGACAATGTGATCTCCCGGATGCTGCAGCCACGGAGGAGCCGAGGACCTTCCAGACGTTCCG
CAAGGCAGAGGCACCAAGAGCTGAGCCCAACAGGCACAGGCTGGCCAGCACCTTTGTCTCGGAGATGAGCTCACT
GCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCGTGGAGGCGCGCTCCAGGCGCTGCGGCGGCTCTCGGT
CTGCGGGAGGACCTCAGTTTAGACTTGGCCACCACTGTCAGCCTCAGGCATGAAAGTGCAAGGGGACCAGGTGG
AAAGACGGGACTGAGGGCAAGAGCAGAGGCAGCAGCAGCAGCAGCAGCAGGTGCTGTGAACATACCTCAGACGCCT
CTGGATCCAAGAACCAGGGGCTGAGGATCTGTGGAACAAGAGCTGGTTTCTAAAATCTTGTAACTCACTAGCTAG
CGGCGGCTGAGAATTTAGGGTGAATGATGCTACCCCCACAGAGGAGGCAAGAGCCCCAGGACTAACAGCTGAC
TGACCAAAGCAGCCCCCTTGTAAAGCAGCTCTGAGTCTTTTGGAGGACAGGGACGGTTTGTGGCTGAGATAAGTGT
TCCTGGCAAACATATGTGGAGCACAAAGGGTCAGTCTCTGGCAGAACAGATGCCACGGAGTATCACAGGCAGG
AAAGGGTGGCCTTCTTGGGTAGCAGGAGTCAGGGGCTGTACCTTGGGGGTGCCAGGAAATGCTCTCTGACCTAT
CAATAAAGGAAAAGCAGTAAAAA

FIGURE 171

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331

<subunit 1 of 1, 1184 aa, 1 stop

<MW: 129022, pI: 5.20, NX(S/T): 5

MMQLLQLLLGLLGPGGYLFLLGDCQEVTTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQ
VLDINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNLTHTYTLSPSEHFALDVIV
GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSNDNSPAFAESS
LALAIQEDAAPGTLIIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGPNIPIAHCKVLIKVLVDVNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQD
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNYEEMAGFEFQVIAEDSGQPMCLASSVSVWVSLDDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTTPPLATHSSRPFLTTT
IVARDADSGANGEPLYRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCQLQAPFHLTPTLYRTLNRNQNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLP
QPATGQPRSRPLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAFAERNPVEELTVDSPPVQQISQLLSLLHQGFQPKPNHRGNKYLAKEPGGS
RSAIPD TDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPD
PAWMARLSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSDLATSAAASGMKVQGDGPGGKTGTGEGKSRGSS
SSSRCL

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 172

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGG
CAGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAG
GCAGGAGCCTTCCTTACACTTCGCCATGAGTTTCTTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTTTTTGGATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTAT
GAGATACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGCAACCAT
GTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACT
GGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTTCATGGTGCCTTTTTTACATTGGC
TATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTTCTGTCTCTT
ATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCCCATTCTCAGCCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAGTGACTCTC
ATGGCTCTTCTTTCTGGATTTGGTGCTGTCAACTGCCCATACACTTACATGTCTTACTTCTT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA
TGATCATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAA
GTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGG
AAGTGAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGC
TTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTC
AAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGGAAAATTTT
CATGGCTACCATCAATATGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCACAAGAGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCCAACACATT
TCCTTCATTCTTGTTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC
CAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCTGCTATTAGCAC
AGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAA
TACCGCACCATAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTT
TGATGTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAAC
AGGCACCAGAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGT
GGTTTCAAATTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAAC
AAACAAAATGCTATGGTAGCATTTTTTACCTTCATAGCATACTCCTTCCCCGTGAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGGAGAACTAACTCAAGACAATACTCA
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGAGCCAAGAACTAA
AGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT
AGGATTTCCGTTTTAAGGTTACATGGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFASFCTMFELIIFEI
LGVLNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFFMYFF
WKLGDPPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQ
QEVDALEELSRQLFLETADLYATKERIEYSKTFKGGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDVPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISS
SKSSNVIVLLLAQIMGYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 175

GTGTTGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCCTAAAATTTGGCCAAGGGTTTC
TTTNTTGAATTCCGGGTTNNGNATACCTTCCCAGAAAATATTTTTTGGATTTGGGGTAGNTT
TTTTTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGGTGATNTT
NTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCATNTTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATC
CTGGTTTTTCATGGTGCCTTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTGCTTTTTTCTGTCTNTTATGGCTGACCTTTATGTATTTNTTNTGGAAAN
TAGGAGATCCCTTTCCCATTC

FIGURE 175

FIGURE 176

CTCGCGCAGGGATCGTCCCATGCCCGGGGCTCGGAGCCGCGACCCCTTGGGGGGCCTCCGGGATTTGCTACCTTTT
TGGCTCCCTGCTCGTCAACTGCTCTTCTCACGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCCTTGCGCAA
GGAGGGCGAGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCAGAGCTG
GCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGGGCAGCAGGCGAATCGCACTGGAGGCCCTCTTCGCTTG
CCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGAAAAGGAAAGCAA
GGAGAACCAGTGGTTGGGAGTCAGTGTTCGGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCACACCCGATA
TGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTGCCTGCTTTGTGCTCAGCCAGGA
CCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGAAGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATT
TGGGTTCTGCCAGCAGGGCACAGCTGCCGCCCTTCTCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAAC
CTATAATTGGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGGACGACGG
TCCCTACGAGGCGGGGGGAGAGAAGGAGCAGGACCCCCGCTCATCCCGGTCCCTGCCAACAGCTACTTTGGCTT
CTCTATTGACTCGGGGAAAGGTCTGGTGCCTGCAGAAAGAGCTGAGCTTTGTGGCTGGAGCCCCCGCGCCAACCA
CAAGGGTGTGTGGTTCATCTGCGCAAGGACAGCGCCAGTCCGCTGGTGCCCGAGGTTATGCTGTCTGGGGAGCG
CCTGACCTCCGGCTTTGGCTACTCAGTGGCTGTGGCTGACCTCAACAGTGATGGCTGGCCAGACCTGATAGTGGG
TGCCCCCTACTTCTTTGAGCGCCAAAGAGAGCTGGGGGGTGTGTGTATGTGTACTTGAACCAGGGGGGTCACTG
GGCTGGGATCTCCCCCTCTCCGGCTCTGCGGCTCCCCTGACTCCATGTTTCGGGATCAGCCTGGCTGTCTGGGGGA
CCTCAACCAAGATGGCTTTCCAGATATTCAGTGGGTGCCCCCTTTGATGGTGATGGGAAAGTCTTCATCTACCA
TGGGAGCAGCCTGGGGGTTGTGCGCCAAACCTTCACAGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTTCGG
CTACTCCCTGTCAAGCAGCTTGGATATGGATGGGAACCAATACCCTGACCTGCTGGTGGGCTCCCTGGCTGACAC
CGAGTGTCTTTCAGGGCCAGACCCATCCTCCATGTCTCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCT
GGAGCAGCCCCAAGTGTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGGTCTGTTTTAGCTACATTGCAGTCCC
CAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGT
TCCCCGTGTGACGTTTCTGAGCCGTAACCTGGAAGAACCAAGCACCAGGCCTCGGGCACCGTGTGGCTGAAGCA
CCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTCGGGCCATTGT
AGTGACCTTGTCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTCCTGGCCAGGGGCTGCCTCCAGTGGC
CCCCATCTCAATGCCACCAGCCAGCACCAGCGGGCAGAGATCCACTTCTGAAGCAAGGCTGTGGTGAAGA
CAAGATCTGCCAGAGCAATCTGACGCTGGTCCACGCCCGCTTCTGTACCCGGGTGAGCGACACGGAATTCACAC
TCTGCCCATGGATGTGGATGGAAACAGCCCTGTTTGCACTGAGTGGGCAGCCAGTCATTGGCCTGGAGCTGAT
GGTCACCAACCTGCCATCGGACCCAGCCAGCCCAAGGCTGATGGGGATGATGCCCATGAAGCCCAGCTCCTGGT
CATGCTTCTGACTCACTGCACCTACTCAGGGGTCCGGGCCCTGGACCCCTGCGGAGAAGCCACTCTGCCTGTCCAA
TGAGAATGCCCTCCCATGTTGAGTGTGAGCTGGGGAAACCCCATGAAGAGAGGTGCCAGGTACCTTCTACCTCAT
CCTTAGCACCTCCGGGATCAGCATTGAGACCAGGAACCTGGAGGTAGAGCTGCTGTTGGCCAGCATGTGGTGAAGA
GGAGTGTGATCCAGTCTCTGCAAGAGCCCGTGTCTTCAATTGAGCTGCCACTGTCCATTGCAGGAATGGCCATTCC
CCAGCAACTCTTCTTCTCTGGTGTGGTGGAGGGGCGAGAGGCCATGCAGTCTGAGCGGGATGTGGGCAGCAAGGT
CAAGTATGAGGTACGGTTTCCAACCAAGGCCAGTCGCTCAGAACCCTGGGCTCTGCCTTCTCAACATCATGTG
GCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGCAGGGGCTGG
GCAGAAAGGGCTTTGCTCTCCAGGCCCAACATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGCGGGA
GCTGGAGCCACCTGAGCAGCAGGAGCCTGGTGAGCGGCAGGAGCCAGCATGTCTGTTGGCCAGTGTCTCTGTC
TGAGAAGAAGAAAAACATCACCCCTGGACTGCGCCCCGGGGCACGGCCAACCTGTGTGGTGTTCAGCTGCCACTCTA
CAGCTTTGACCGCGCGGCTGTGCTGCATGTCTGGGGCCGTCTCTGGAACAGCACCTTTCTGGAGGAGTACTCAGC
TGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCAGAGTGAAGTCTCCATAAAGAACTTGATGCTCCGAGA
TGCCTCCACAGTGATCCCACTGATGGTATACTTGGACCCCATGGCTGTGGTGGCAGAAGGAGTGCCCTGGTGGGT
CATCCTCCTGGCTGTACTGGCTGGGCTGTGGTGTCTAGCACTGTGGTGTCTCCTGTGGAAGATGGGATTCTT
CAAACGGGGCAAGCACCCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAAGATTCTCGGGAAGACCGACAGCA
GTTCAAGGAGGAGAAGACGGGCACCATCCTGAGGAACAACCTGGGGCAGCCCCCGCGGGAGGGCCCGGATGCACA
CCCCATCCTGGCTGTGACGGGCATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCTAGGTTCC
CATGTCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGT
GGGCTGTGGTGTGCGATCAAGATTGGCAGGATCGGCTTCTCAGGGGCAAGACCTCTCCACCCACAAGAAC
TCCTCCCACCAACTTCCCCTTAGAGTGTCTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGG
TGAGAAGGGCAGGGGTGCTCTGATGCAAAGGTGGGGAGAAGGGATCCTAATCCCTTCTCTCCATTACCCCTGT
GTAACAGGACCCCAAGGACCTGCCCTCCCCGGAAGTGCCTTAACCTAGAGGGTCGGGGAGGAGGTTGTGTCACTGA
CTCAGGCTGTCTTCTCTAGTTTCCCCTCTCATCTGACCTTAGTTTGTGCCATCAGTCTAGTGGTTTCGTGGT
TTCGTCTATTTATTAATAAATAATTTGAGAACAAAAA

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWL
GVSVRSQGPGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDDGGEWKFC
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGRTARVELCAQGSADLAHLDDGPYEA
GGEKEQDPRLIPVPANSYFGFSIDSGKGLVRAEELSFVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPDLLIVGAPYFFERQEELGGAVYVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDNLQDGFDPDIAVGAPFDGDGKVFIYHGSSLGVVAKPSQVLE
GEAVGIKSFGYSLSGSLDMDGNQYPDLLVGLADTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGHGSVCVDLRVCFYSIAVPSSYSPTVALDYVLDADTDRLRGQVPRVTFLSRNLEPK
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSYSLQTPRLRRQAPGQGLPPVAP
ILNAHQPSTQRAEIHFLKQCGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTTALFA
LSGQPVIGLELMVTNLPSDPAQPQADGDDAHEAQLLVMLPDSLHYSVGRALDPAEKPLCLSN
ENASHVECELGNPMKRGAQVTFFYLILSTSGISIIETTELEVELLLATISEQELHPVSARARVF
IELPLSIAGMAIPQQQLFFSGVVRGERAMQSERDVGSKVKEYEVTFSNQGQSLRTLGS AFLNIM
WPHEIANGKWL LYPMQVELEGGQGPGQKGLCSRPNHLHLDVDSRDRRRRELEPPEQQEPGE
RQEPSMSWWPVSSAEKKKNITLDCARGTANCVVFCPLYSFDRAAVLHVWGRLWNSTFLEEY
SAVKSLEVIVRANITVKSSIKNLMLRDASTVIPVMVYLDPMVVAEGVPWWVILLAVLAGLL
VLALLVLLLWKMGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNNWGS PRREGP
DAHPIAADGHPELGPDPGHPGPGTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCTGGCT
CACAAACAAGATGCTCAAGGTGTCAGCCGTA CTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCA
GTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGGGCGGTCCGACGGCGGTAAATTTTC
TGGATGATAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCCGACAGTGGAAC
AAATTC CGAGACGAAGTAGAGGATGATTATTTCCGCACCTTGAGTCCAGGAAAACCTTCGA
TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTG
CCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTCAGATGGTCATACCTACTCTTTTCAGTGCA
AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCTTCAGATAAGCCCACAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGA CTGGTTC AAGGCCCTTCATGAAAAGTGGAA
GTCAAAACAAGAAGACAAAAACATTGCTGAGGCCTGAGAGAAGCAGATTCGATACCAGCATC
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAAACAGACTTGATACAAACTATGACCT
GCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAATGAACAGTGTACCAAGG
CATTCCTCAATTCTTGTGACACATACAAGGACAGTTTAATATCTAATAATGAGTGGTGCTAC
TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA
AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCATGGCAGTGTGGACAGTGCTGGTGTGTTGACAGATATGGAAATGAAGTC
ATGGGATCCAGAATAAATGGTGTGTCAGATTGTGCTATAGATTTTGAGATCTCCGAGATTT
TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGTATGAC
CATGATGTATACATTTGATTTGATGACAGTTGAAATCAATAAATCTACATTTCTAATATTTA
CAAAAATGATAGCCTATTTAAAATTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCA
CATATATTTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAAT
AAGAATCATTTGCTTTGAGTTTTTATATTCCTTACACAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTTCACGAGAACAACTTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAG
ATAATTCTAAGTGAAATTTAAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGG
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAAAAAA

FIGURE 179

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829

><subunit 1 of 1, 436 aa, 1 stop

><MW: 49429, pI: 4.80, NX(S/T): 0

MLKVSAVLCVCAAAWCSQSLAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCP
SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKTKTLLRPERSRFDTSILPI
CKDSLGMFMNRLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ
RQQDPPCQTELSNIQKRQGVKKLLGQYIPLCEDGDGYYKPTQCHGSGVQCWCVDTRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDEDEDDIMNDEDEIEDDDEDEGDDDDGGDDHDVYI

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 180

CAGACTCCAGATTTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGC
 CTCTTTAGCCCCGGGATCGCCCCAGCAGGGATGGGCGACAAGATCTGGCTGCCCTTCCCCGTGCTCCTTCTGGCC
 GCTCTGCCTCCGGTGCTGCTGCTGGGGCGGGCGGCTTACACCTTCCCTCGATAGCGACTTACCTTTACCTTT
 CCCGCCGGCCAGAAGGAGTGCTTCTACCAGCCCATGCCCTGAAGGCTCGCTGGAGATCGAGTACCAAGTTTGA
 GATGGAGCAGGATTAGATATTGATTTCCATCTTGCCCTCTCCAGAAGGCAAAACCTTAGTTTTTGAACAAAGAAAA
 TCAGATGGAGTTTCACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTTCAGCACCATT
 TCTGAGAAGGTGATTTTCTTTGAATTAATCCTGGATAATATGGGAGAACAGGCACAAGAACAAGAAGATTGGAAG
 AAATATATTACTGGCACAGATATATTGGATATGAAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC
 AGACTAAGCAAAAAGTGGGCACATACAAATTTCTGCTTAGAGCATTTGAAGCTCGTGATCGAAACATACAAGAAAGC
 AACTTTGATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTTCATGGTGGTGGTGTGAGCCATTCAAGTTTAT
 ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAAGTAGAACTTAAACTCCAACTAGAGTACGTAACATTGAAA
 AATGAGGCATAAAAATGCAATAAACTGTTACAGTCAAGACCATTAAATGGTCTTCTCCAAAATATTTTGAGATATA
 AAAGTAGGAAAACAGGTATAATTTAATGTGAAAATTAAGTCTTCACTTTCTGTGCAAGTAATCCTGCTGATCCAG
 TTGTACTTAAGTGTGTAACAGGAATATTTTGCAATATAGGTTTAACTGAATGAAGCCATATTAATAACTGCAT
 TTTCTTAATTTGAAAAATTTTGCAATGTCTTAGGTGATTTAAATAAATGAGTATTGGGCCATAATTGCAACACC
 AGTCTGTTTTTAACAGGTTCTATTACCCAGAACTTTTTTGTAAATGCGGCAGTTACAAATTAAGTGTGGAAGTTT
 TCAGTTTTAAGTTATAAATCACCTGAGAATTACCTAATGATGGATTGAATAAATCTTTAGACTACAAAAGCCCAA
 CTTTTCTCTATTACATATGCATCTCTCCTATAATGTAAATAGAATAATAGCTTTGAAATACAATTAGGTTTTTG
 AGATTTTTATAACCAATACATTTTCACTGTAAATATAGCAGAAAGCATTAGTCTTTGTACTTTGCTTACATTC
 CAAAAGCTGACATTTTACGATTCTTAAAAACACAAAGTTACACTTACTAAAATTAGGACATGTTTCTCTTTG
 AAATGAAGAATATAGTTTAAAGCTTCTCCTCCATAGGGACACATTTTCTCTAACCTTAACTAAAAGTGTAGGA
 TTTTAAAAATTAAGTGTGAGGTAAAATAAGTTTATTTTAAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA
 TAATCATGTTATGTTAATTTTAAATGATTGCTGACTTGGATAATTCAATTATTACCAGCAGTTATGAAGGAAATA
 TTGCTAAAATGATCTGGGCTTACCATAAATAAATATCTCCTTTTCTGAGCTCTAAGAAATATCAGAAAAACAGGAA
 AGAATTTAGAAAAACTTGAAGAAACCTAATCCAAAATAAATTCACCTAAGTAGAATATAAATAAATATCTAGA
 ATCTGACTGGCTCATCATGACATCCTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
 AAATTTGGCTGTAGGTTTTTATTTTCTACAAGAATTCTGGTTTGAATTATTTTGTAAAGCAGGTACATTTTATA
 AAATGTAAAGCCTTGGGAGAGGTTTAGCACTGGGTGTACATTTTATTTAAAAATTTTATTATAACAACCTTTTAT
 TAAAATGGCCTTTCTGAACACTTTATTTATTGATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTTTAAA
 CACCTAAATGTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTGAAGTCTATGGGGGTCTTAC
 TCAAGTACTAGTAATTTAACTTCATCATGAATGAACATAAATTTTAAAGTTATGCCCATTTATAACGTTGTTTAT
 GACTACATTTGTGAGTTAGAAAACAACTTAAATTTGGGGTATAGAACCCCTCAACAGGTTAGTAATGCTGGAAT
 CTTGATGAGCAATAATGATAACCAGAGAGTGATTTTCACTTACACTCATAGTAGTATAAAAAGAGATACATTTCCC
 TCTTAGGCCCTTGGGAGAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAAAATGAGGTAAATGCCGTAT
 ATGATCAATTACCTTAAATTTGGCCAAGAAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCAACACTTGCAGAACAA
 AGGTCAATAAGATCCTTGCTTATGAATACCCCTCCCTTTTGGCTGTTAAATTTGCAATGAGAAGCAAAATTTACA
 GTACCATAACTAATAAAGCAGGGTACAGATATAAACTACTGCATCTTTTCTATAAACTGTGATTAAGAAATCTA
 CCTCTCCTGTATGGCTGTTACTGTACTGTACTCTCTGACTCCTTACCTAACAAATGAATTTGTTACATAATCTTCT
 ACATGTATGATTTGTGCCACTGATCTTAAACCTATGATTCAGTAACCTTCTTACCATATAAAAACGATAATTTGCTT
 TATTTGGAAAAGAAATTTAGGAATACTAAGGACAATTATTTTATAGACAAAGTAAAAAGACAGATATTTAAGAGG
 CATAACCAAAAAAGCAAACTTGTAAACAGAGTAAAAATCTTTAATATTTCTAAAGACATACTGTTTATCTGCTT
 CATATGCTTTTTTTAATTTCACTATTCATTTCTAAATTAAGTTATGCTAAATTTAGTAAGCTGTTTATCACTT
 AACAGCTCATTTTGTCTTTTCAATATACAAATTTTAAAAATACTACAATATTTAACTAAGGCCAACCGATTTT
 CATAATGTAGCAGTTACCGTGTTCACCTCACACTAAGGCCTAGAGTTTGGCTCTGATATGCATTTGGATGATTAAT
 GTTATGCTGTTCTTTTCACTGTGAATGTCAAGACATGGAGGGTGTGTAATTTTATGGTAAAATTAATCCTTCTTA
 CACATAATGGTGTCTTAAATTTGACAAAAATGAGCACTTACAATTGTATGTCTCCTCAATGAAGATTCTTTAT
 GTGAAATTTTAAAGACATTGATTCGCATGTAAGGATTTTTCATCTGAAGTACAATAATGCACAATCAGTGTTG
 CTCAACTGCTTTTATCTTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGTATATAATAA
 AATTTATCAAAGGAAAA

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196

><subunit 1 of 1, 229 aa, 1 stop

><MW: 26017, pI: 4.73, NX(S/T): 0

MGDKIWLPFPVLLLAALPPVLLPGAAGFTPSLDSDFTLTPAGQKECFYQPMPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCFDNTFSTISEKVIFFEL
ILDNMGEQAQEQEDWKKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

FIGURE 182

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAGAT
CTCACCAGAGAGTTCGCAGACACTATGCTGCCTCCCATGGCCCTGCCAGTGTGTCCTGGATG
CTGCTTTCCTGCCTCATTCTCCTGTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAACTGCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTTGTACCAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACTCTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC
TCCACCATCTTAAACCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCTATGTCTGCAAGTTCAAGGACTAGGGCAGGT
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCTGGAAGAGAATATTCTCCCCAACTGCCCTACCTGACTACCTTGTCATGATCCTCC
TTCTTTTTTCCTTTTTCTTCACCTTCATTTAGGCTTTTCTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 183

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965

<subunit 1 of 1, 175 aa, 1 stop

<MW: 19330, pI: 7.25, NX(S/T): 1

MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM

DADLACQKRPSGKLVSVLGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSS

TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTTCTCCCTCTT
GAGTCCTTCTGAGATGATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTGCGGATGG
TAGCGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACCTCGGTTCTC
AATTCCAACGCTATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTC
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTT
CTGATCAAAATCATTTCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGTAAACCTGTCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTC
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTGAGA
GACACTAAACCAGCTATCCAAATGCAGTGAACTCCTTTTATATAATAGATGCTATGAAAACC
TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCAT
TCCAATAACACCTTCCAAAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACCTCCCTG
TGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCAATGA
AACTTTTAATTATTTTTCTAAAGGTGCTGCACTGCCTATTTTTCTCTTGTATGTAAATTT
TTGTACACATTGATTGTTATCTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCAT
TCAGCTTATAGTTCTTAAAAGCATAACCTTTACCCCATTTAATTCTAGAGTCTAGAACGCA
AGGATCTCTTGGAATGACAAATGATAGGTACCTAAAATGTAACATGAAAATACTAGCTTAT
TTCTGAAATGTACTATCTTAATGCTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGA
AATAAAATTTAACATTTAAAAA

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA
APGILYPGGNKYQTTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRKRCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQ
KDDHQASNSSRLHTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

FIGURE 188

TGTGTTTCCCTGCAGTCAGAATTTGGGACNGCAGGGGTTCCTGGACCTGATTTTGCAGCGGA
ACGGGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTTTCTTCTCCTTCNG
GAGTCCTTNTGAGANGATGGTTTTTGGGCGCAGCGGGAGCTAACCCGGTTTTTTGTNGCGATG
GTAGCGGCGGTTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTTC
AATTCCAACGNTATCAAGAACCTGCCCCACCGNTGGGCGGCGCTGCGGGGCACCCAGGNTT
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGGANGCGGGCGTGCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCANGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTNTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGG

FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCGCTGGTAGTCGCCGGTGTGGCTGCACCTCACCAATCCCGTGCGCCGCGG
CTGGGCCGTGCGAGAGTGCCTGTCTCTCTGACGCGGTGCTTGGGCTCGGCCAGGCGGGGTCCGCCGCCA
GGGTTTGGAGATGGGGGAGTAGCTACAGGAAGCGACCCCGCGATGGCAAGGTATATTTTGTGGAATGAAAAGGA
AGTATTAGAAATGAGCTGAAGACCATTACAGATTAATATTTTGGGGACAGATTTGTGATGCTTGATTACCCCT
TGAAGTAATGTAGACAGAAGTTCTCAAATTTGCATATTACATCAACTGGAACCAGCAGTGAATCTTAATGTTAC
TTAAATCAGAACTTGCATAAGAAAAGAGAATGGGAGTCTGGTTAAATAAAGATGACTATATCAGAGACTTGAAAAG
GATCATTCTCTGTTTTCTGATAGTGTATATGGCCATTTTAGTGGGCACAGATCAGGATTTTACAGTTTACTTGG
AGTGTCCAAAACCTGCAAGCAGTAGAGAAATAAGACAAGCTTTCAAGAAATTGGCATTGAAGTTACATCCTGATAA
AAACCCGAATAACCCAAATGCACATGGCGATTTTAAATAAATAAGAGCATATGAAGTACTCAAAGATGAAGA
TCTACGGAAAAAGTATGACAAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGGCCAGTATGAAAGCTGGAA
CTATTATCGTTATGATTTTGGTATTTATGATGATGATCCTGAAATCATAACATTGGAAAGAAGAGAATTTGATGC
TGCTGTTAATTTCTGGAGAACTGTGGTTTGTAAATTTTACTCCCAGGCTGTTACACTGCCATGATTTAGCTCC
CACATGGAGAGACTTTGCTAAAGAGTGGATGGGTTACTTCGAATTGGAGCTGTTAACTGTGGTGTATAGAAAT
GCTTTGCCGAATGAAAGGAGTCAACAGCTATCCAGTCTCTTCAATTTTTCGGTCTGGAATGGCCCGCAGTGAATA
TCATGGAGACAGATCAAAGGAGAGTTTAGTGAGTTTTCGAATGCAGCATGTTAGAAGTACAGTGACAGAACTTTG
GACAGGAAATTTTGTCAACTCCATACAAACTGCTTTTGTGCTGGTATTGGCTGGCTGATCACTTTTGTTCAAA
AGGAGGAGATTTGTTGACTTCACAGACACGACTCAGGCTTAGTGGCATGTTGTTTCTCAACTCATTGGATGCTAA
AGAAATATATTTGGAAGTAATACATAATCTTCCAGATTTTGAAGTACTTTCCGGCAAACACACTAGAGGATCGTTT
GGCTCATCATCGGTGGCTGTTATTTTTCATTTTGGAAAAAATGAAAATTCAAATGATCCTGAGCTGAAAAAAT
AAAACTCTACTTAAAAATGATCATATTCAAGTTGGCAGGTTTGACTGTTTCTCTGCACCAGACATCTGTAGTAA
TCTGTATGTTTTTTCAGCCGTCTCTAGCAGTATTTAAAGGACAAGGAACCAAAGAATATGAAATTCATGTAATA
GAAGATTTCTATATGATATATCTTGCTTTTGCCAAAGAAAGTGTGAATTTCTCATGTTACCACGCTTGGACCTCAAAA
TTTTCTTGCCAAATGACAAAGAACCATGGCTTGTGATTTCTTTGCCCCCTGGTGTCCACCATGTCGAGCTTTACT
ACCAGAGTTACGAAGAGCATCAAATCTTCTTTATGGTCAGCTTAAGTTTGGTACACTAGATTGTACAGTTTCATGA
GGGACTCTGTAACATGTATAACATTCAGGCTTATCCAAACAACAGTGGTATTCAACCAGTCCAACATTCATGAGTA
TGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTCATAGAGGATCTTATGAATCCTTCAGTGGTCTCCCTTAC
ACCCACCACCTTCAACGAACCTAGTTACACAAAGAAAACACAACGAAGTCTGGATGGTTGATTTCTATTCTCCGTG
GTGTCTATCCTTGCCCAAGTCTTAAAGCAGAAATGGCCCGGACATTAAGTGGACTGATCAACGTGGG
CAGTATAGATTGCCAACAGTATCATTTCTTTTGTGCCAGGAAAACGTTCAAAGATACCCTGAGATAAGATTTTT
TCCCCCAAATCAAATAAAGCTTATCAGTATCAGATTACAATGGTTGGAATAGGGATGCTTATCCCTGAGAAT
CTGGGGTCTAGGATTTTACCTCAAGTATCCACAGATCTAACACCTCAGACTTTCAGTGAAGAAAGTTCTACAAGG
GAAAAATCATTGGGTGATTGATTTCTATGCTCCTTGGTGTGGACCTTGCCAGAATTTTGCTCCAGAATTTGCTCAGCATGCCAGAA
CTTGGCTAGGATGATTAAAGGAAAAGTGAAGCTTGGAAAAGTAGACTGTGAGGCTTATGCTCAGCATGCCAGAA
AGCTGGGATCAGGGCTTATCCAAGTCTTAAAGTTTATTTCTACGAAAGAGCAAAGAGAAATTTTCAAGAAGAGCA
GATAAATACCAGAGATGCAAAAGCAATCGCTGCCCTTAATAAGTGAAGAAATTTGGAACTCTCCGAAATCAAGGCAA
GAGGAATAAGGATGAACTTTGAATAATGTTGAAGATGAAGAAAAAGTTTAAAGAAATTTCTGACAGATGACATCAG
AAGACACCTATTTAGAATGTTACATTTATGATGGGAATGAATGAACATTATCTTAGACTTGCAGTTGTACTGCCA
GAATTATCTACAGCACTGGTGTAAAAGAAGGGTCTGCAAACTTTTTCTGTAAAGGGCCGGTTTATAAATATTTTA
GACTTTGCAGGCTATAATATATGTTTCACACATGAGAACAAGAATAGAGTCATCATGTATTCTTTGTTATTTGCT
TTTAACAACCTTTAAAAAATATTAACAGATTCTTAGCTCAGAGCCATACAAAAGTAGGCTGGATTCACTCCATG
GACCATAGATTGCTGTCCCCCTCGACGGACTTATAATGTTTCAGGTGGCTGGCTTGAACATGAGTCTGCTGTGCT
ATCTACATAAATGTCTAAGTTGTATAAAGTCCACTTTCCCTTCACGTTTTTTGGCTGACCTGAAAAGAGGTAAT
TAGTTTTTGGTCACTTGTTCTCTTAAAAATGCTATCCCTAACCATATATTTATATTTCTGTTTTAAAAACACCCAT
GATGTGGCACAGTAAACAAACCTGTTATGCTGTATTATATGAGGAGATTCTTCATTGTTTTCTTTCTTCTCA
AAGTTTGAAGAAATGCTTTTAAATTTTTCACAGCCGAGAAACAGTGCAGCAGTATATGTGCACACAGTAAGTACAC
AAATTTGAGCAACAGTAAGTGCACAAATCTGTAGTTTGTGTATCATCCAGGAAAACCTGAGGGAAGAAAAATTA
TAGCAATTAAGTGGCATTGTAGAGTATCCTAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATA
TGTGTTTCATGATTTTCTGAAATGCTTTTCATAGAAATTTCCCACTGATAGTTGATTTTGGAGGCATCTAATAT
TTACATATTTGCTTCTGAACCTTTGTTTTGACCTGTATCCTTTATTTACATTGGGTTTTTCTTTTCATAGTTTGG
TTTTTCACTCCTGTCCAGTCTATTTATTTTCAAATAGGAAAAATTAATTTACAGGTTGTTTTACTGTAGCTTAT
AATGATACTGTAGTTATTTCCAGTTACTAGTTTACTGTGAGAGGCTGCCTTTTTCAGATAAATATTGACATAATA
ACTGAAGTTATTTTATAAGAAATCAAGTATATAAATCTAGGAAAGGATCTTCTAGTTTCTGTGTTGTTTGA
CTCAAAGAATCACAATTTGTCAGTAACATGTAGTTGTTAGTTTATAATTCAGAGTGTACAGAATGGTAAAAAT
CCAATCAGTCAAAGAGGTCAATGAATTAAGGCTTGCACCTTTTCAAAAAAAAAAAAAAAAAA

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439

<subunit 1 of 1, 747 aa, 1 stop

<MW: 86127, pI: 7.46, NX(S/T): 2

MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNPNNAHGDFLKNRAYEVLKDEDLRKKYDKYGEKGLEDNQGGQYESWNYRYDFGI
YDDDPEIITLERREFDAAVNSGELWFVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS
IQTAFAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHRWLLFFHFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSPLYVFQP
SLAVFKGQGTKEYEIHGKKILYDILAFAKESVNSHVTTLGPNFPANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVFVNQSNIEYEGHHS
AEQILEFIEDLMNPSVVSLLPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLG
FLPQVSTDLTPTQTFSEKVLQGKNHWVIDFYAPWCGPCQNFAPFELLARMIKGKVKAGKVDC
QAYAQTCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAI AALISEKLETLRNQGKRKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAA
GCCATGAACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGA
GTCGTTGGTGAAGTTTTTTCATTCCTCAGAGGAGAAAATCTGTGGCTGGGGAGATTGTTCTCA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTTGCAAAACGACAGAGC
ATATTGGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAACTGCAGCTGAGTGCCGAAA
ACTAGGCGTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA
CATCCTAGGACATTTTTTGGATCACAAAAGCACTTCTTCCATCGATGATGGAGAGAAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCCCTACCTCATCCCATAT
TGTTCCAGCAAATTTGCCGCTGTTGGCTTTCACAGAGGTCTGACATCAGAACTTCAGGCCTT
GGGAAAACTGGTATCAAAACCTCATGTCTCTGCCAGTTTTTGTGAATACTGGGTTCACCA
AAAATCCAAGCACAAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTTTTGTTCATCGTATATCAATATCTTTCTGAG
ACTACAGAAGTTTCTTCTGAACGCGCCTCAGCGATTTTAAATCGTATGCAGAATATTCAAT
TTGAAGCAGTGGTTGGCCACAAAATCAAAATGAAATGGAATAAATAAGCTCCAGCCAGAGATG
TATGCATGATAATGATATGAATAGTTTTGAATCAATGCTGCAAAGCTTTATTTACATTTTT
TCAGTCCTGATAATATTA AAAACATTGGTTTGGCACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGTCTTCTGTTTTCTCAAGAATATTTACGTAGTTTTTTCATAGGTCTGTTTTTCTT
TCATGCCTCTTAAAACTTCTGTGCTTACATAAACATACTTAAAAGGTTTTCTTTAAGATAT
TTTATTTTTTCCATTTAAAGGTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAAGAGAACT
TATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTCCAATCTGTAGCCATGCCACAG
AATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTT
ATCTCAACCTGGACATATTTTAAGATTGAGCATTGAAAGATTTCCCTAGCCTCTTCTTTT
TCATTAGCCCAAAACGGTGCAACTCTATTCTGGACTTTATTACTTGATTCTGTCTTCTGTAT
AACTCTGAAGTCCACCAAAAGTGGACCCTCTATATTTCTCCCTTTTTATAGTCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTTTAAATCTCAGAATTTTAAGTTCTAGCCCCATGA
TAACCTTTTTCTTTGTAATTTATGCTTTCATATATCCTTGGTCCCAGAGATGTTTAGACAAT
TTTAGGCTCAAAAATTAAAGCTAACACAGGAAAAGGAACTGTACTGGCTATTACATAAGAAA
CAATGGACCCAAGAGAAGAA

FIGURE 192

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

<subunit 1 of 1, 300 aa, 1 stop

<MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSI
LVLWDINKRGVEETAAECKRLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDTVIVVNNAGTV
YPADLLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYC
SSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFNVTGFTKNPSTRLWPVLETDEVVRSLLID
GILTNNKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCCAGCAGG
ATGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGGAGAGGGCCCAGCCCCGCGGGGC
AGGATGACCAAGCCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTTCATGATCCT
GCTGATCATCGTGTACTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGCACACGTCTTCT
CTAGGCCGCACACGGGGCCGCGCTGCCCACGCCCGGGCCGGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTTCGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGA
CCTTCCCAGAAAGGAGACGGAGCAGCCGCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAG
GCTACGACTGGTCCCCGCGCGACGCCCGGCGCAGCCAGACCAGGGCCGGCAGCAGGCGGAG
CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCAACCTCCAGCCTGGCCTTCCCCACCAAGGAGCG
CGCATTTCGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGACGACCGGCACGGGG
CCATCTACTGCTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCA
CGTGACAACGCCAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCCTCTTCGTGCGCGACCCC
TTCGTGCGCCTGATCTCCGCCTTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTCGCCGTGCCCATGCTGCGGCTGTACGCCAACCACACCAGCCTGCCCGCCTCGGCGC
GCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCTTCGCCAACTTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCA
CCCGTGCCAGATCGACTACGACTTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGC
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCGAGCTACCGGAACAGG
ACCGCCAGCAGCTGGGAGGAGGACTGGTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCC
GAGACTTGAAAGCTTTTCGCGTTGCTTTTTTCTCGCGTGCCCTGGAACCTGACGCACGCGCACTCC
AGTTTTTTTATGACCTACGATTTTGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATATTGTTTTTTAAGATTAATATATTTTCAGGTATTTAATACGA

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112

<subunit 1 of 1, 414 aa, 1 stop

<MW: 48414, pI: 9.54, NX(S/T): 4

MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFYLHTSF SRPHTGPPLPTPGPDRDRELT
DSDVDEF LDKFLSAGVKQSDLP RKETE QPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFTP KERA FDDIPNSEL SHLIVDDRHGAIYCYVPKVACTNWK RVMIVLS
GSL LHRGAPYRDPLRIPREHVHNASAH LTFNKFWRRY GKL SRHLMKVKLKKYTKFLFVRDPF
VRLISAFRSKF ELENEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLET LDEDA AQLLQLLQVDRQLRFPPSYRNRT
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLF GYPK PENLLRD

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

FIGURE 194

FIGURE 195

TCGGGCCAGAATTTCGGCACGAGGCGGCACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGA
AAGAGGCCCCAGAGTAGAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACGCGCT
ATGCCGGGAAGGTGGTGGTTCGTGACCGGGGGCGGGCGCGGCATCGGAGCTGGGATCGTGCGC
GCCTTCGTGAACAGCGGGGCCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGCCGGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCCTGGTTTTCTGAGACCATCCGCCGATTTGGCCGCCTGGATTGTGTTGTCAACAAC
GCTGGCCACCACCCACCCACACAGAGGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCCTACCTGCGGA
AGAGTCAAGGGAATGTCATCAACATCTCCAGCCTGGTGGGGGCAATCGGCCAGGCCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGACCAAAGCTTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCCAGGAAACATCTGGACCCCGCTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCTAGGGCCACAATCCGAGAGGGCATGCTGGCCCAG
CCTTGGGCCCGCATGGGCCAGCCCGCTGAGGTGCGGGCTGCGGCAGTGTTCCTGGCCTCCGA
AGCCAACTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGGTGCAGAGCTGGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCGTGGACGCCCCGATATCCCTTCCTTGATTTCTCTCATTT
CTACTTGGGGCCCCCTTCCTAGGACTCTCCCACCCCAAACCTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAAGCCCAGTTAGCAAGGTGCCGGGTACCCTGCAGGTTCCCAT
AAAAACGATTTGCAGCC

FIGURE 196

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045

<subunit 1 of 1, 270 aa, 1 stop

<MW: 28317, pI: 6.00, NX(S/T): 1

MATGTRYAGKVVVVTGGGRGIGAGIVRAFVN SGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGR LDCVVNNAGHHPPPQRPEETSAQGFRQLLELNLLGTYTTLTKL
ALPYLRKSQGNVINIS SLVGAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGN
IWTPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTTGG
AELGYGCKASRSTPVDAPDIPS

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGAATGGACTGGCCTCACAACCTG
CTGTTTCTTCTTACCATTTCCATCTTCCTGGGGCTGGGCCAGCCAGGAGCCCCAAAAGCAA
GAGGAAGGGGCAAGGGCGGCCTGGGCCCCTGGCCCCTGGCCCTCACCAGGTGCCACTGGACC
TGGTGTACGGATGAAACCGTATGCCCCGATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGTGAACCCCTTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTTCAGCCAGGTTCTGTGCGCCGCCG
CCTCTGCCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG
CTGTGGGCTGCACCTGCATCTTCTTGAATCACCTGGCCCAGAAGCCAGGCCAGCAGCCCGAGA
CCATCCTCCTTGACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAA
GCAAG

FIGURE 198

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294

<subunit 1 of 1, 180 aa, 1 stop

<MW: 20437, pI: 9.58, NX(S/T): 1

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEEY
ERNIEEMVAQLRNSSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPRIIPVDLPEARCLCL
GCVNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPPRTGPCRQRAVMETIavgctcIF

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

FIGURE 199

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCCGGCGGG
CGAGCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCGCTCG
GCGCCCAACATGGCGGGTGGGCGCTGCGGGCCGCGAGCTAACGGCGCTCCTGGCCGCCTGGAT
CGCGGCTGTGGCGGCGACGGCAGGCCCGAGGAGGCCGCGCTGCCGCCGGAGCAGAGCCGGG
TCCAGCCCATGACCGCCTCCAACCTGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTT
TACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATTGAGAAATGGGAGGCTTTTGCAAAGAA
TGGTGAAATACTTCAGATCAGTGTGGGGAAGGTAGATGTGATTCAAGAACAGGTTTGAGTG
GCCGCTTCTTTGTCAACACTCTCCAGCATTTTTTTCATGCAAAGGATGGGATATTCGCCCGT
TATCGTGGCCCAGGAATCTTCGAAGACCTGCAGAAATTATATCTTAGAGAAGAAATGGCAATC
AGTCGAGCCTCTGACTGGCTGGAAATCCCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTC
TTTTTTAGCATCTCTGGCAAGATATGGCATCTTCACAACTATTTTACAGTGACTCTTGGAATT
CCTGCTTGGTGTCTTATGTGTTTTTCGTGATAGCCACCTTGGTTTTTGGCCTTTTTTATGGG
TCTGGTCTTGGTGGTAATATCAGAAATGTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGC
GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCCGGAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAACAAAGACAGCCTTGAGATGATGAAGAAGA
GAAAGAAGATCTTGGCGATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACTTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGCCCCCAGGAGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCCTGCCCAGC
TGACACAGAGGTGGTGGAAGACTCCTTGAGGCAGCGTAAAAGTCAGCATGCTGACAAGGGAC
TGTAGATTAAATGATGCGTTTTTCAAGAATACACACCAAACAATATGTGAGCTTCCCTTTGG
CCTGCAGTTTGTACCAAATCCTTAATTTTTTCTGAATGAGCAAGCTTCTCTTAAAGATGCT
CTCTAGTCATTTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGT
GACAATCAGGATATAGAAAAACAAACGTAGTGTGGGATCTGTTTGGAGACTGGGATGGGAA
CAAGTTCATTTACTTAGGGGTGAGAGAGTCTCGACCAGAGGAGGCCATTCCAGTCCTAATC
AGCACCTTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCT
CCTGAGCATCCCCAAAGTGTAACGTAGAAGCCTTGCATCCTTTTCTTGTGTAAAGTATTTAT
TTTTGTCAAATTGCAGGAAACATCAGGCACCACAGTGCATGAAAAATCTTTCACAGCTAGAA
ATTGAAAGGGCCTTGGGTATAGAGAGCAGCTCAGAAGTCATCCAGCCCTCTGAATCTCCTG
TGCTATGTTTTATTTCTTACCTTTAATTTTTTCCAGCATTTCCACCATGGGCATTCAGGCTCT
CCACACTCTTCACTATTATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC
TGTGTTTTGTTCAATTCTGACCTAAGGGGTTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
GACTGCCAAACATCTCAAATGAAATGTTGTGGCCATCAGAGACTCAAAGGAAGTAAGGATT
TTACAAGACAGATTAAAAAAAATTGTTTTGTCCAAAATATAGTTGTTGTTGATTTTTTTTTT
AAGTTTTCTAAGCAATATTTTTTCAAGCCAGAAGTCCTCTAAGTCTTGCCAGTACAAGGTAGT
CTTGTGAAGAAAAGTTGAATACTGTTTTGTTTTTCTCTCAAGGGGTTCCCTGGGTCTTGAAC
TACTTTAATAATAACTAAAAAACCACTTCTGATTTTCTTTCAGTGATGTGCTTTTGGTGAAA
GAATTAATGAACTCCAGTACCTGAAAGTGAAAGATTTGATTTTGTTCATCTTCTGTAAATC
TTCCAAAGAATTATATCTTGTAAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTTCTTT

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433

<subunit 1 of 1, 349 aa, 1 stop

<MW: 38952, pI: 4.34, NX(S/T): 1

MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAP
WCPSCQQTDSEWEAFKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGI FRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISSECFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEEK
DSDNEEENKDSLVDDEEEKEDLGDEDEAE EEEEEEDNLAAGVDEERSEANDQGPPGEDGV TRE
EVEPEEAEEGISEQPCPADTEVVEDSLRQRKRSQHADKGL

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

FIGURE 201

ATCTGGTTGAACTACTTAAGCTTAATTTGTTAAACTCCGGTAAGTACCTAGCCCACATGATT
TGA CTCAGAGATTCTCTTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAGAAAAGAGCTCC
CAAATGCTATATCTATT CAGGGGCTCTCAAGAACAAATGGAATATCATCCTGATTTAGAAAAT
TTGGATGAAGATGGATATACTCAATTACACTTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTTCAGAGAAAAGGATCGTGTGCTGCATCTCCTCCTTGGCGCCTCATTGCTGTAATTTTGG
GAATCCTATGCTTGGTAATACTGGTGATAGCTGTGGTCCTGGGTACCATGGGGGTTCTTTCC
AGCCCTTGTCCTCCTAATTGGATTATATATGAGAAAGAGCTGTTATCTATT CAGCATGTCACT
AAATTCCTGGGATGGAAGTAAAAGACAATGCTGGCAACTGGGCTCTAATCTCCTAAAGATAG
ACAGCTCAAATGAATTGGGATTTATAGTAAAACAAGTGTCTTCCCAACCTGATAATTCATTT
TGGATAGGCCTTTCTCGGCCCCAGACTGAGGTACCATGGCTCTGGGAGGATGGATCAACATT
CTCTTCTAACTTATTT CAGATCAGAACACAGCTACCCAAGAAAACCCATCTCCAAATTGTG
TATGGATT CACGTGTCAGTCATTTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTTGT
GAGAAGAAGTTTTCAATGTAAGAGGAAGGGTGGAGAAGGAGAGAGAAATATGTGAGGTAGTA
AGGAGGACAGAAAACAGAACAGAAAAGAGTAACAGCTGAGGTCAAGATAAATGCAGAAAATG
TTTAGAGAGCTTGGCCAACTGTAATCTTAACCAAGAAATTGAAGGGAGAGGCTGTGATTTCT
GTATTTGTGACCTACAGGTAGGCTAGTATTATTTTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCCAAGCTTGAGTTTTTATTTTTTATTTATTTTATTTTTTTGAGATAGGGTCT
CACTTTGTTACCCAGGCTGGAGTGCAGTGGCACAATCTCGACTCACTGCAGCTATCTCTCGC
CTCAGCCCCCTCAAGTAGCTGGGACTACAGGTGCATGCCACCATGCCAGGCTAATTTTTTGGTG
TTTTTTGTAGAGACTGGGTTTTGCCATGTTGACCAAGCTGGTCTCTAACTCCTGGGCTTAAG
TGATCTGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGCCACCACACCTGGC
CCCAAGCTTGAATTTTTCATTTCTGCCATTGACTTGGCATTACCTTGGGTAAGCCATAAGCGA
ATCTTAATTTCTGGCTCTATCAGAGTTGTTTCATGCTCAACAATGCCATTGAAGTGCACGGT
GTGTTGCCACGATTTGACCCTCAACTTCTAGCAGTATATCAGTTATGAACTGAGGGTGAAAT
ATATTTCTGAATAGCTAAATGAAGAAATGGGAAAAAATCTTCACCACAGTCAGAGCAATTTT
ATTATTTTCATCAGTATGATCATAATTATGATTATCATCTTAGTAAAAAGCAGGAACTCCTA
CTTTTTCTTTATCAATTAAATAGCTCAGAGAGTACATCTGCCATATCTCTAATAGAATCTTT
TTTTTTTTTTTTTTTTTTTTTTTGAGACAGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAACGG
CACGATCTCGGCTCACCGCAACCTCCGCCCCCTGGGTTCAAGCAATCTCCTGCCTCAGCCT
CCCAAGTAGCTGGGATTACAGTCAGGCACCACCACACCCGGCTAATTTTGTATTTTTTTTAGT
AGAGACAGGGTTTTCTCCATGTCGGTCAGGGTAGTCCCGAACTCCTGACCTCAAGTGATCTGC
CTGCCTCGGCCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT
TGTATAATATGTAATTGTAGGGAACTGCTCTCATAGGAAAGTTTTCTGCTTTTTTAAATACA
AAAATACATAAAAAATACATAAAATCTGATGATGAATATAAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAACATTTTGAATATGTTTTATTAGTTTTGTGATGTACTGTTTTACAATTTTT
ACCATTTTTTTT CAGTAATTACTGTAAAATGGTATTATTGGAATGAACTATATTTCTCATG
TGCTGATTTGTCTTATTTTTTTT CATACTTTCCCACTGGTGCTATTTTTTATTTCCAATGGATA
TTTCTGTATTACTAGGGAGGCATTTACAGTCCTCTAATGTTGATTAATATGTGAAAAGAAAT
TGTACCAATTTTACTAAATTATGCAGTTTAAAATGGATGATTTTATGTTATGTGGATTT CAT
TTCAATAAAAAAAACTCTTATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912

<subunit 1 of 1, 201 aa, 1 stop

<MW: 22563, pI: 4.87, NX(S/T): 1

MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSAAASPWRLLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNWIIYEKSCYLFSMSLNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

FIGURE 203

GGAAGGGGAGGAGCAGGCCACACAGGCCACAGGCCGGTGAGGGACCTGCCAGACCTGGAGGGTCTCGCTCTGTCA
CACAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCATCGTAACCTCCACCTCCCGGGTTCAAGTGATTCTCATGCC
TCAGCCTCCCGAGTAGCTGGGATTACAGGTGGTGACTTCCAAGAGTGACTCCGTCCGGAGGAAAATGACTCCCCAG
TCGCTGCTGCAGACGACACTGTTCTGCTGAGTCTGCTCTTCTGGTCCAAGGTGCCACGGCAGGGGCCACAGG
GAAGACTTTTCGCTTCTGCAGCCAGCGGAACCAGACACACAGGAGCAGCCTCCACTACAAACCCACACCAGACCTG
CGCATCTCCATCGAGAACTCCGAAGAGGCCCTCACAGTCCATGCCCCCTTCCCTGCAGCCCACCTGCTTCCCGA
TCCTTCCCTGACCCAGGGGCCCTTACCACCTTCTGCCTCTACTGGAACCGACATGCTGGGAGATTACATCTTCTC
TATGGCAAGCGTGACTTCTTGTGAGTGACAAAGCCTCTAGCCTCCTCTGCTTCCAGCACCAGGAGGAGAGCCTG
GCTCAGGGCCCCCGCTGTTAGCCACTTCTGTACCTCCTTGGTGGAGCCCTCAGAACATCAGCCTGCCAGTGCC
GCCAGCTTACCTTCTCCTTCCACAGTCTTCCCCACACGGCCGCTCACAATGCCTCGGTGGACATGTGCGAGCTC
AAAAGGGACCTCCAGCTGCTCAGCCAGTTCTGAAGCATCCCCAGAAGGCCTCAAGGAGGCCCTCGGCTGCCCCC
GCCAGCCAGCAGTTGCAGAGCCTGGAGTCGAAACTGACCTCTGTGAGATTTCATGGGGGACATGGTGTCTTTCGAG
GAGGACCGGATCAACGCCACGGTGTGGAAGCTCCAGCCCACAGCCGCCCTCCAGGACCTGCACATCCACTCCCGG
CAGGAGGAGGAGCAGAGCGAGATCATGGAGTACTCGGTGCTGCTGCCCTCGAACACTCTTCCAGAGGACGAAAGGC
CGGAGCGGGGAGGCTGAGAAGAGACTCTCTGTTGGACTTCAGCAGCCAAGCCCTGTTCCAGGACAAGAATTCC
AGCCAAGTCTTGGGTGAGAAGGTCTTGGGGATTGTGTTACAGAACACCAAAGTAGCCAACCTCACGGAGCCCGTG
GTGCTCACTTTCCAGCACCAGCTACAGCCGAAGAATGTGACTCTGCAATGTGTGTTCTGGGTTGAAGACCCACA
TTGAGCAGCCCCGGGCATTGGAGCAGTGTGGGTGTGAGACCGTCAGGAGAGAAACCCAAACATCCTGCTTCTGC
AACCACCTTGACCTACTTTGCAGTGTGATGGTCTCCTCGGTGGAGGTGGACCCGTGCACAAGCACTACCTGAGC
CTCCTCTCCTACGTGGGCTGTGTGCTCTTGCCTTGGCCTGCCTTGTACCATTGCGCCTACCTCTGCTCCAGG
GTGCCCCCTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCACATGAACCTGCTGCTGGCCGTCTT
CTGCTGGACACGAGCTTCTGCTCAGCGAGCCGGTGGCCCTGACAGGCTCTGAGGCTGGCTGCCGAGCCAGTGCC
ATCTTCTGCACTTCTCCTGCTCACCTGCCTTCTCTGGATGGGCCTCGAGGGGTACAACCTCTACCGACTCGTG
GTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTT
CTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACATATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAG
GGCGTCATCTACCTTCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATACCAACCTGGGCCTCTTCAGC
CTGGTGTCTTCTGTTCAACATGGCCATGCTAGCCACCATGGTGGTGACATCTTGCCTGCGCCCTGCGCCCCACACCCAA
AAGTGGTCACATGTGCTGACACTGCTGGCCCTCAGCCTGGTCTCTTGGCCTGCCCTGGGCCTTGATCTTCTTCTCC
TTTGCTTCTGGCACCTTCCAGCTTGTGCTCCTCTACCTTTTTCAGCATCATCACCTCCTTCCAAGGCTTCTCATC
TTCATCTGGTACTGGTCCATGCGGCTGCAGGCCCGGGGTGGCCCTCCCTCTGAAGAGCAACTCAGACAGCGCC
AGGCTCCCCATCAGCTCGGGCAGCACCTCGTCCAGCCGCATCTAGGCCTCCAGCCCACCTGCCCATGTGATGAAG
CAGAGATGCGGCCTCGTCGCACACTGCCTGTGGCCCCCGAGCCAGGCCCAGCCCCAGGCCAGTCAGCCGCAGACT
TTGGAAGCCCCAACGACCATGGAGAGATGGGCCGTTGCCATGGTGGACGGAATCCCGGGCTGGGCTTTTGAATTG
GCCTTGGGGACTACTCGGCTCTCACTCAGCTCCACCGGACTCAGAAGTGCGCCGCTATGCTGCCTAGGGTACTG
TCCCCACATCTGTCCCAACCCAGCTGGAGGCCTGGTCTCTCCTTACAACCCCTGGGCCCAGCCCTCATGTCTGGG
GGCCAGGCCTTGGATCTTGAAGGTCTGGCACATCCTTAATCCTGTGCCCTGCCTGGGACAGAAATGTGGCTCCA
GTTGCTCTGTCTCTCGTGGTCACCTGAGGGCACTCTGCATCCTCTGTCAATTTTAACTCAGGTGGCACCCAGGG
CGAATGGGGCCCAGGGCAGACCTTCAGGGCCAGAGCCCTGGCGGAGGAGAGGCCCTTTGCCAGGAGCACAGCAGC
AGCTCGCCTACCTCTGAGCCCAGGCCCTCCTCCTCAGCCCCCAGTCCCTCCCTCCATCTTCCCTGGGGTTCT
TCCTCCTCTCCCAGGGCCTCCTTGTCTCCTTCTGTTCAAGCTGGGGGTCCCCGATTCCAATGCTGTTTTTTGGGGA
GTGGTTTTCCAGGAGCTGCCTGGTGTCTGCTGTAAATGTTTGTCTACTGCACAAGCCTCGGCCTGCCCTGAGCCA
GGCTCGGTACCGATGCGTGGGCTGGGCTAGGTCCCTCTGTCCATCTGGGCCTTTGTATGAGCTGCATTGCCCTTG
CTCACCTGACCAAGCACACGCCTCAGAGGGGCCCTCAGCCTCTCCTGAAGCCCTCTTGTGGCAAGAACTGTGGA
CCATGCCAGTCCCGTCTGGTTTTCCATCCCACCACTCCAAGGACTGAGACTGACCTCCTCTGGTGACACTGGCCTA
GAGCCTGACACTCTCCTAAGAGGTTCTCTCAAGCCCCCAAATAGCTCCAGGCGCCCTCGGCCGCCCATCATGGT
TAATTCGTGTCCAACAAACACACACGGGTAGATTGCTGGCCTGTTGTAGGTGGTAGGGACACAGATGACCGACCTG
GTCACCTCCTCCTGCCAACATTAGTCTGGTATGTGAGGCGTGCCTGAAGCAAGAACTCCTGGAGCTACAGGGACA
GGGAGCCATCATTCTGCCTGGGAATCCTGGAAGACTTCTGCAGGAGTCAGCGTTCAATCTTGACCTTGAAGAT
GGGAAGGATGTTCTTTTTACGTACCAATTCTTTTGTCTTTTGATATTAAAAAGAAATACATGTTTCATTGTAGAGA
ATTTGGAAGCTGTAGAAGAGAATCAAGAAGAAAAATAAAAAATCAGCTGTTGTAATCGCCTAGCAAAAAAAAAAAAA
AA

FIGURE 204

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921

<subunit 1 of 1, 693 aa, 1 stop

<MW: 77738, pI: 8.87, NX(S/T): 7

MTPQSLLQTTFLLSLLFLVQGAHGRGHREDFRCSQRNQTHRSSLHYKPTPDLRISIENSE
EALT VHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLLSDKASSLLCFQH
QEESLAQGPPLLATSVTSWWSPQNISLPSAASF TFSFHSPHTAAHNASVDMCELKRDLQLL
SQFLKHPQKASRRPSAAPASQQQLQSLESKLT SVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEEQSEIMEYSVLLPRTL FQRTKGRSGEAEKRLLL VDFSSQALFQDKNSSQVLGE
KVLGIVVQNTK VANLTEPVVLT FQHQLQPKNVT LQCVFWVEDPTLSSPGHWSSAGCETVRRE
TQTSCFCNHLTYFAVLMVSSVEVDAVHKHYLSLLSYVGC VVSALACLVTIAAYLCSRVP LPC
RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAI FLHFSLLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVDNYGPI ILAVHRTPEGVIY
PSMCWIRDSLVSYITNLGLFSLVFLFNMA MLATMVVQILRLRPHTQKWSHVL TLLGLSLVLG
LPWALIFFSFASGTFQLVVL YLFSIITSFQGF LIFIWYWSMRLQARGGPSPLKSNSDSARLP
ISSGSTSSSRI

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

FIGURE 205

TGCCTGGCCTGCCTTGTCAACAATGCCGCTTACTCTGCTTCCAGGTTGCCCTGCCTTGCAGA
GGAAANCNTCGGGACTACACCNTCAAGTGCACATGAACCTGCTGCTGGCCGTCCTTCCTGCTG
GACACGAGCTTCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTTCCTGCACCTTCTCCTGCTCACCTGCCTTTCCTGGATGGGCCTCGAGGGG
TACAACCTCTACCGACTCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAA
GCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTTCTGGTGACGCTGGTGGCCCTGGTGGATG
TGGACAACTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCCCT
TCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCAACCTGGGCCTCTTCAGCCT
GGTGTTTCTGTTCAACATGG

FIGURE 206

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTTCAGGTCCAGGTTTTGCTTTGA
TCCTTTTCAAAAACCTGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTTTGGATGGGATTATGTGGAAACTACCCCT
GCGATTCTCTGTGTCAGAGCAGGCTCGGCGCTTCCACCCAGTGCAGCCCTCCCTGGCGGTGGTGAAAGAGAC
TCGGGAGTCCGCTGCTTCCAAAGTGCCCGCCGTGAGTGAGCTCTCAGGCGGAAATCCAACCTGAGTAGTAAATTCC
AGTTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATTACTGTGTCTACTAATG
GAAGTATTCAGAGCCCAAGGTTTCTCATACTTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
AGGAAAATGTATGGATACAACCTTACGTTTGATGAAAGATTTGGGCTTGAAGACCCAGAAGATGACATATGCAAGT
ATGATTTTGTAGAAGTTGAGGAACCCAGTGATGGAACATATATTAGGGCGCTGGTGTGGTTCTGGTACTGTACCAG
GAAAACAGATTTCTAAAGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCTTCTGAACCAGGGT
TCTGCATCCACTACAACATTGTCTATGCCACAATTCACAGAAGCTGTGAGTCTTCAGTGCTACCCCTTCAGCTT
TGCCACTGGACCTGCTTAATAATGCTATAACTGCCTTTAGTACCTTGAAGACCTTATTCGATATCTTGAACCAG
AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGGCTTTTGTTTTGGAA
GAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCT
CAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATAACCTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTG
GTGGGAACGTGTCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACC
ACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTGAGGGGATTGCACAAATCACTACCCGACGTGGCCCTGGAGC
ACCATGAGGAGTGTGACTGTGTGTGTCAGAGGGAGCACAGGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCA
GAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCT
TCAAGGACCTTTTCATCTTCAGGATTTACAGTGCATTTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCA
ACAGCTCTTTTGGAGAGGAGGCTAAAGGACAGGAGAGAAAGGTCCTTCAATCGTGGAAAGAAAATTAATGTTGTAT
TAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCTAGCTGGGTTCTGTATTTTCAGTTCTTTC
GATACGGCTTAGGGTAATGTGAGTACAGGAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAAC
TCTAAAGCTCCATGTCTGGGCCTAAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTTGCTCATATTCACAT
ATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAACCTTGT
GTCATGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAATAATTTCTGCCATTTAGAAGAAGAGAACTACA
TTCATGGTTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTG
TTTTATTGTGTACATTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCT
ATTTTTACCAAAGGTATTTAATATTCTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCT
AAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATGTTGCTCTGACAAAAATACATGTATTTCA
TTCTCGTATGGTGTAGAGTTAGATTAACTGTCATTTTAAAAAAGTGAATTTGGAATAGAATTGGTAAGTTGCAA
GACTTTTTGAAAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGA
AAGTAGACATTAGATCCAGCCATTACTAACCTATTCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACAT
AAAGCACCTTGAAAAAGACTTGGCAGCTTCCGTGATAAAGCGTGCTGTGCTGTGCAGTAGGAACACATCCTATTTA
TTGTGATGTTGTGGTTTTTATTATCTTAAACTCTGTTCCATACACTGTATAAATACATGGATATTTTTATGTACA
GAAGTATGTCTCTTAACAGTTCACTTATTGTACTCTGGCAATTTAAAAAGAAAATCAGTAAATATTTTTGCTTGT
AAAATGCTTAATATNGTGCCTAGGTTATGTGGTGACTATTTGAATCAAAAATGTATTGAATCATCAAATAAAGA
ATGTGGCTATTTTGGGGAGAAAATTAATAAAAAAAAAAAAAAAAAAAGGTTTAGGGATAACAGGGTAATGCGGCC

FIGURE 207

MSLFGLLLLSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPKGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLEEVRLY
SCTPRNFSVSIREELKRDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSKVTKKYHEVLQ
LRPKTGVRGLHKS LTDVALEHHEECDVCVRGSTGG

Signal sequence:

amino acids 1-14

FIGURE 208

CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCCATTTTGGGAAGA
AGACTAAAAATGGTGTTTCCAATGTGGACACTGAAGAGACAAATCTTATCCTTTTAAACATAATCCTAATTTCC
AAACTCCTTGGGGCTAGATGGTTTTCTAAACCTCTGCCCTGTGATGTCACCTCTGGATGTTCCAAAGAACCATGTG
ATCGTGGACTGCACAGACAAGCATTGTACAGAAATTCCTGGAGGTATCCCACGAACACCACGAACCTCACCCCTC
ACCATTAAACCACATACCAGACATCTCCCCAGCGTCCTTTCACAGACTGGACCATCTGGTAGAGATCGATTTTCAGA
TGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAAACCCAGAAGC
TTTAGTGGACTCACTTATTTAAATCCCTTTACCTGGATGGAAACCAGCTACTAGAGATACCGCAGGGCCTCCCCG
CCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACACATCTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC
AACATAGAAATACTCTACCTGGGCCAAAACCTGTTATTTATCGAAATCCTTGTATGTTTCATATTCAATAGAGAAA
GATGCCTTCTTAACTTGACAAAGTTAAAGTGCTCTCCCTGAAAGATAACAATGTACAGCCGTCCCTACTGTT
TTGCCATCTACTTTAACAGAACTATATCTCTACAACAACATGATTGCAAAAATCCAAGAAGATGATTTTAATAAC
CTCAACCAATTACAAATCTTGGACCTAAGTGGAAATGCCCCCTGTTTATAATGCCCATTTCTTGTGCGCCG
TGTAATAATAATTCTCCCCTACAGATCCCTGTAAATGCTTTTGATGCGCTGACAGAATTAAGGTTTTACGTCTA
CACAGTAACTCTCTCAGCATGTGCCCCAAGATGGTTTAAAGACATCAACAACTCCAGGAACCTGGATCTGTCC
CAAACTTCTTGGCCAAAGAAATGGGGATGCTAAATTTCTGCATTTTCTCCCCAGCCTCATCCAATTGGATCTG
TCTTTCAATTTTGAACCTCAGGTCTATCGTGCATCTATGAATCTATCACAAGCATTTTCTTCACTGAAAAGCCTG
AAAATTCTGCGGATCAGAGGATATGCTTTTAAAGAGTTGAAAAGCTTTAACCTCTCGCCATTACATAATCTTCAA
AATCTTGAAGTTCTTGATCTTGGCACTAACTTTTATAAAATTGCTAACCTCAGCATGTTTAAACAATTTAAAGA
CTGAAAGTCATAGATCTTTCAGTGAATAAAATATCACCTTCAGGAGATTCAGTGAAGTTGGCTTCTGCTCAAAT
GCCAGAACTTCTGTAGAAAGTTATGAACCCAGGTCCCTGGAACAATTACATTATTTTCAATATGATAAGTATGCA
AGGAGTTGCAGATTCAAAAACAAGAGGCTTCTTTCATGTCTGTTTAAAGAGCTGCTACAAGTATGGGCAGACC
TTGGATCTAAGTAAAAATAGTATATTTTGTCTCAAGTCTCTGATTTTTCAGCATCTTTCTTCTCAAATGCCTG
AATCTGTGAGAAATCTCATTAGCCAACTCTTAATGGCAGTGAATTCACCTTTAGCAGAGCTGAGATATTTG
GACTTCTCCAACAACCGGCTTGATTTACTCCATTCAACAGCATTTGAAGAGCTTCACAACTGGAAGTTCTGGAT
ATAAGCAGTAATAGCCATTATTTTCAATCAGAAGGAATTACTCATATGCTAAACTTTACCAAGAACCTAAAGGTT
CTGCAGAACTGATGATGAACGACAATGACATCTCTCTCCACCAGCAGGACCATGGAGAGTGAGTCTCTTAGA
ACTCTGGAATTCAGAGGAAATCACCTTAGATGTTTATGAGAGAGAAGGTGATAACAGATACTTACAATTATTCAAG
AATCTGCTAAATTAGAGGAATTAGACATCTCTAAAAATTCCTAAGTTTCTTGCCTTCTGGAGTTTTTGTATGGT
ATGCCTCCAAATCTAAAGAATCTCTCTTTGGCCAAAAATGGGCTCAAATCTTTCAGTTGGAAGAACTCCAGTGT
CTAAAGAACCTGGAACTTTTGGACCTCAGCCACAACCAACTGACCACTGTCCCTGAGAGATTATCCAAGTGTTC
AGAAGCCTCAAGAATCTGATTTCTTAAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCCTTC
CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCCAGAAAATGTCCTC
AACAATCTGAAGATGTTGCTTTTGCATCATAATCGGTTTCTGTGCACCTGTGATGCTGTGTGGTTTTGTCTGGTGG
GTTAACCATACGGAGGTGACTATTCCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGC
CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATA
TCTGTATCTCTCTTCTCATGGTGATGATGACAGCAAGTCACTCTATTTCTGGGATGTGTGGTATATTTACCAT
TTCTGTAAGGCCAAGATAAAGGGGTATCAGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTTATTGTGTAT
GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAACTGGAAGACCCAAGAGAGAAA
CATTTTAATTTATGCTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTCTGGAAAACCTTTCCAGAGCATA
CAGCTTAGCAAAAAGACAGTGTGTTGTGATGACAGACAAGTATGCAAAAGACTGAAAATTTAAGATAGCATTTTAC
TTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAGCCCTTTCAGAAG
TCCAAGTTCTCCAGCTCCGGAAGGCTCTGTGGGAGTTCTGTCTTGAGTGGCCAACAAACCCGCAAGCTCAC
CCATACTTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGAA
ACGGTCTAGCCCTTCTTTGCAAAACACAACCTGCCTAGTTTACCAAGGAGAGGCCTGGC

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGG
IPTNTTNLTLTINHIPDISPASFHRLDHLVEIDFRNCNCPVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLGQNCYR
NPCYVSYSEKDAFLNLTKLKVLSLKDNNVTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPWRW
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSLKSL
KILRIRGYVFKEKLSFNLSPLHNLQNLVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKIS
PSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQ
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDIFSNNRLDLLH
STAFEELHKLEVLDISSNSHYFQSEGITHMLNFTKNLKVLQKLMMNDNDISSSTSRTMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNSLSFLPSGVFDGMPPNLKNLSL
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTDAVWFVWVWNHTEVTIP
YLATDVTCVGPGAHKGQSVISLDLYTCELDLTNLILFSLISISVSLFLMVMMTASHLYFWDVW
YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPAVTEWVLAELVAKLEDPREKHFNLCLLE
RDWLPGPVLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVIILIFLE
KPFQKSKFLQLRKRLCGSSVLEWPTNPQAHFYPFWQCLKNALATDNHVAYSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

[illegible]

GGGTACCAATTCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACACAGAAACATGGAAAAATGTTCTCTTCA
AGTCGTCAATGCTGACCTGCATTTTCTGCTAATATCTGGTTCTGTGAGTTATGCGCCGAAGAAAAATTTTCTA
GAAGCTATCCTTGTGATGAGAAAAAGCAAATGACTCAGTTATTGCAGAGTGCAGCAATCGTCGACTACAGGAAG
TTCCCCAAACGCTGGGCAAATATGTGACAGAACTAGACCTGTCTGATAAATTCATCACACACATAACGAATGAAT
CATTTCAAGGGCTGCAAAATCTCACTAAAAATAAATCAAACCAACCAATGTAACAGCACAGGAAATC
CCGGTATACAATCAAAATGGCTTGAATATCACAGCGGGCATTCTCAACCTAGTAACACCTAAGGGAGTTACTGC
TTGAAGACCAACCACTTACCCCAAATACCTCTGGTTTGCCAGAGTCTTTGACAGAACTTAGTCTAATTCAAACA
ATATATACAACATAACTAAAGAGGGCATTTCAAGACTTATAAACTTGAAAAATCTCTATTTGGCCTGGAAGTCT
ATTTTAAACAAAGTTTGCGAGAAAACTAACATAGAAGATGGAGTATTTGAAACGCTGACAAATTTGGAGTTGCTAT
CACTATCTTTCAATTTCTCTTTCACACGTTGCCACCCAAACTGCCAAGCTCCCTACGCAACTTTTTCTGAGCAACA
CCCAGATCAAAATCACTTAGTGAAGAAGATTTCAAGGCTTGATAAATTAACATTAGATTAAAGCGGGAAC
GTCCGAGGTGCTTCAATGCCCCATTTCCATGCGTGCCTTGTGATGGTGGTGCTTCAATTAATATAGATCGTTTTG
CTTTTCAAACCTTGACCCAACTTCGATACCTAAACCTCTCTAGCACTTCCCTCAGGAAGATTAATGCTGCCTGGT
TTAAAAATATGCCTCATCTGAAGGTGCTGGATCTTGAATCAACTATTTAGTGGGAGAAATAGTCTCTGGGGCAT
TTTTAAACGATGCTGCCCCGCTTAGAAATACCTTGACTTGTCTTTTAACTATATAAAGGGGAGTTATCCACAGCATA
TTAATATTTCCAGAAACTTCTCTAAACTTTTTGTCTCTACGGGCATTTGCATTTAAGAGGTTATGTTTCCAGGAAC
TCAGAGAAGATGATTTCCAGCCCCCTGATGCAGCTTCCAACTTATCGACTATCAACTTGGGTATTAATTTTATTA
AGCAAAATCGATTTCAAACTTTTCCAAAAATTTCCAACTCGGAAATTATTTACTTGTGAGAAAAAGAAATATCAC
CGTTGGTAAAAGATACCCGGCAGAGTTATGCAAATAGTTTCTCTTTCAACGTCATATCCGAAACGACGCTCAA
CAGATTTTGGAGTTTGACCCACATTCGAACTTTTATCATTTCAACCGTCTTTAATAAAGCCACAATGTGCTGCTT
ATGGAAAAAGCCTTAGATTTTAGCCCTCAACAGTATTTTCTTCAATTGGGCCAAACCAATTTGAAAATCTTCTGACA
TTGCCCTGTTTTAAATCTGTCTGCAAAATAGCAATGCTCAAGTGTTAAGTGGAACTGAATTTTCCAGCAATCTCTCATG
TCAAATATTTGGATTTTGACAAACAATAGACTAGACTTTGATAAGCTAGTGCTCTTACTGAATTGTCCGACTTGG
AAGTTCCTAGATCTCAGCTATAATTCACACTATTTTCAAGAAATAGCAGGCGTAACACATCATCTAGAATTTATTCAAA
ATTTTCACAAATCTAAAAGTTTTTAAACTTTGAGCCACAACAACATTTTATACTTTAAACAGATAAGTATAACCTGGAAA
GCAAGTCCCTGGTAGAATTAGTTTTTCAAGTGGCAATCGCCTTGACATTTTGTGGAATGATGATGACAAACAGGTATA
TCTCCATTTTCAAAGGTCTCAAGAACTCTGACACGCTGCGATTTATCCCTTAATAGGCTGAAGCACATCCCAAATG
AAGCATTCCTTAATTTGGCAGCGAGTCTCACTGCACTACATATAAGTATAATGTTAAAGTTTTTAACTTGA
CATTAATCTCAGCAGTTTCTCGTCTCGAGTTGCTTGACTTACGTGGAAACAACTACTCTTTTTAACTGATAGCC
TATCTGACTTTTACATCTTCCCTTCGGACACTGCTGCTGAGTCATAACAGGATTTCCACCTACCTCTGGCTTTC
TTTTCTGAAGTCAGTAGTCTGAAGCACCTCGATTTAAGTTCCAATCTGCTAAAAACAATCAACAAATCCGCACTTG
AAACTAAGACCACCACCAAATTTATCTATGTTGGAACCTACCGGAAACCCCTTTGAATGCACCTGTGACATTTGGAG
ATTTCCGAAGATGGATGGATGAACATCTGAATGTCAAAATTTCCAGACTGGTAGATGTCAATTTGTCGAGCTCTG
GGGATCAAGAGGGGAAGAGTATTTGTGAGTCTGGAGCTAAACAATCTGTGTTTCAGATGTCACTGCAGTGATATTAT
TTTTCTTACAGTTCTTTTATCACCACCATGGTTATGTTGGCTGCCCTGGCTCACCATTGTGTTTTACTGGGATGTTT
GGTTTTATATATAATGTGTGTTTTAGCTAAGGTAAAAGGCTACAGGTCTCTTTCCACATCCCAAACCTTCTATGATG
CTTACATTTCTTATGACACCAAGATGCCTCTGTTACTGACTGGGTGATAAATGAGCTGCGCTACCACCTTGAAG
AGAGCCGAGACAAAAACGTTCTCTTTGCTAGAGGAGAGGGATTTGGGACCCGGGATTTGGCCATCATCGACAACC
TCATGCGAGAGCATCAACCAAGACAAGAAAAACAGTATTTGTTTTAAACCAAAAAATATGCAAAAACTGGAACCTTTA
AAACAGCTTTTTTACTTGGCTTTGAGAGGCTAATGGATGAGAACATGGATGTGATTATATTTATCCTGCTGGAGC
CAGTGTACAGCATCTCAGTATTTGAGGCTACGGCAGCGGATCTGTAAGAGCTCCATCCTCCAGTGGCCTGACA
ACCCGAAGGCAGAAAGGCTTGTGTTTTGGCAAACCTCTGAGAAATGTGGTCTTGACTGAAAAATGATTACCGGTATAACA
ATATGTATGTCGATTTCCATTAAGCAATACTAACTGACGTTAAGTCATGATTTTCGCGCCATAATAAAGATGCAAG
GAATGACATTTCTGATTTAGTTATCTATTGCTATGTAACCAATATCCCAAACTTAGTGGTTTTAAACACACACA
TTTGTGCTGGCCCAAGTTTTTGAAGGTCAGGAGTCAGGACCCAGTAACTGGGTCCTCTGCTCAGGGTGTCTCAG
AGGCTGCAATGTAGGTGTTTACCAGAGACATAGGCATCAGTGGGTCACACTCATGTGGTTGTTTTCTGGATTCA
ATTCCTCCTGGGCTATTGGCCAAAGGCTATACCTCATGTAAGCCATGCGAGCCTCTCCACAAAGGCAGCTTGCTTC
ATCAGAGCTAGCAAAAAAGAGAGGTTGCTAGCAAGATGAAGTCACAATCTTTTGTAAATCGAATCAAAAAAGTGAT
ATCTCATCACTTTGGCCATATTTCTATTGTTAGAAAGTAACCAAGTCCACAGGCTCCATGGAGTGACCAC
TCAGTCCAGGGAACACAGCTGAAGACCAAGATGGTGAGCTCTGATGCTTCAGTTGGTTCATCAAGTATTTTCCCT
TGACTGCTGTCTGGGATGGCCTGCTATCTTGATGATAGATTGTGAATATCAGGAGGCAGGGATCACTGTGGACC
ATCTTAGCAGTTGACCTAACACATCTTCTTTTCAATATCTAAGAACTTTTGGCACTGTGACTAATGGTCTTAATA
TTAAGCTGTTGTTTTATATTTATCATATATCTATGGCTACATGGTTATATTATGCTGTGGTTGCGTTTCGGTTTTAT
TTACAGTTGCTTTTTCAAAATATTTGCTGTAAACATTTGACTTCTAAGGTTTAGATGCCATTTAAGAACTGAGATGG
ATAGCTTTTTAAAGCATCTTTTACTTCTTACCATTTTTTAAAGATGACAGCTAAATTCGAAGCTTTTTGGTCTATA
TTGTTAAATTGCCATTTGCTGTAATCTTAAATGAATGAATAAAAAATGTTTCATTTTACAAAAA

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSPCDEKKQNDSVIAECSNRRLQEVPTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHONGNPGIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTELSQLQNNIYNITKEGISRLINLKNLYLAWNCYFNKV
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL
TLLDLSGNCPRCFNAPFPCVPCDGGASINIDRFQNLTLQLRYLNLSSTSLRKINAAWFKNM
PHLKVLDLEFNILVGEIVSGAFLTMLPRLEILDLSFNIIKGSYPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFIKQIDFKLFQNFNLEIIYLSNRISPLV
KDTRQSYANSSSFQRHIRKRRSTDFEFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACNLNSANSNAQVLSGTEFSAIPHVKYLDLTNNRLDFDNASALTELSDEV
LDLSYNSHYFRIAGVTHHLEFIQNFTNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRL
DILWNDDDNRYISIFKGLKNLTRLDSLNLRLKHIPNEAFLNLPASLTELHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFLTDSLSDFTSSLRTLLLSHNRISHLPSGFLSEVSSLKHLDL
SNLLKTINKSALETKTTTKLSMLELHGNPFECTCDIGDFRRWMDEHLNVKI PRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFFTFFITTMVMLAALAHHLFYWDVWFIYNVCLAKVK
GYRSLSTSQTIFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWDPLAIIID
NMQSINQSKKTVFVLTKKYAKSWNFKTA FYLALQRLMDENMDVIIIFILLEPVLQHSQYLRL
RQRICKSSILQWPDNPKAEGLFWQTLRNVVLTENDSRYNMYVDSIKQY

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

FIGURE 212

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGA
GCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAA
GGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTTGGCAGTGGGCGGCACAGAGCAGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGT
CCGGGCTCACGGGGACCCCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCCTCA
CCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGGCTGGCCCCCTGCCAGGCCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGAC
CAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGA
GCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCA
GATGTGGATGAATGCAGTGCTAGGAGGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGGTGCCAGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGC
CCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCC
TGGTGCACCTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTCCTTCCTG
GAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGCTG
GACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAG
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCC
TTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC
CCCTGGCTACCCCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTG
AGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAG
GCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACATAAAAAATGAAACGTGA
AAAGGGCGGCCGCGACTCTAGAGT
CGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTLCVPGGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

Downloaded from www.sciencedirect.com

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC
AGCAGCATCAGAGCAGCCCCTGTGGTTGGCAGCAAAGTTCAGCTTGGCTGGGCCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGGCCCCGGGGAGGGGAAGTGGCCCCGAGGGAGAGGAACCCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG
CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCTGTCCGGGGGATGACTGATTC
TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGCAGTGGGCGGCACAGAGCACGCCTA
CCGGCCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTG
TGCAGCGTGTGTACCAGCCCTTCTCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTAC
CGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTA
CGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCTGGGGCCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCA
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGC
ATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGAC
TCCCTGAGCGAGCAGATTTCTTCTTCTGGAGGAGCAGCTGGGGTCTTGCTCCTGCAAGAAAGA
CTCGTGACTGCCCAGCGCTCCAGGCTGGACTGAGCCCCCTACGCCGCCCTGCAGCCCCCATG
CCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTTCCTCCTCTTCTCCTCCCCCTTCTCGGGAGGCTCCCCAGACCCTGGCATGGGAT
GGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCA
TCCCAAGGCCAGGTGGACCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGAC
CCATGGCACAGGCCAGGCAGCCCGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGAC
CCCCAGCACAATAAAAATGAAACGTG

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGC
CCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGC
CTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGC
AGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCGTAGGGTGTGTGCTGTCCGGGCTCACG
GGGACCCTGTCTCCGAGTCGTTCGTGCAGCGTGTGTACCAGCCCTTCTCACCACCTGCGAC
GGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGG
GCTGGCCCCCTGCCAGGCCTCGCTACGCGTGTGCCCCGGCTGGAAGAGGACCAGCGGGCTTC
CTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGGGCGGCTGTCCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT
GCCAGTGTGGAGGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGG
CCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACA
GCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCCTCC
TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTCTTCTGGAGGAGCAGCT
GGGGTCTGCTCCTGCAAGAAAGACTCGTGACTGCCAGCGCCCCAGGCTGGACTGAGCCCC
TCACGCCGCCCTGCAGCCCCCATGCCCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCTCCTCTTCTCCTCCCCTTCTCCTCGGGAG
GCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACC
CCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTAC
GAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGG
CCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACATAAAAATGAAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCQSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSLSADGTLCVPGGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 218

GGTTGCCACAGCTGGTTTAGGGCCCCGACCACTGGGGCCCCCTTGTCAGGAGGAGACAGCCTCCCGGCCCGGGGAG
GACAAGTCGCTGCCACCTTTGGCTGCCGACGTGATTCCTGGGACGGTCCGTTTCCTGCCGTACGTGCCGGCCG
AGTTGGGTCTCCGTGTTTTAGGCCGGCTCCCCCTTCCCTGGTCTCCCTTCTCCCGCTGGGCCGGTTTATCGGGAGG
AGATTGTCTTCCAGGGCTAGCAATTGGACTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGAT
TTCAAAGCTGGGCTCAGCCTCTGTTTCTTCTCTCGTGAATCGCAAAACCCATTTTGAGCAGGAATTCCAATCA
TGTCTGTGATGGTGGTGAGAAAGAAGGTGACACGGAAATGGGAGAACTCCAGGCAGGAACACCTTTTGCTGTG
ATGGCCCGCTCATGATGGCCCGGCAAAAGGGCATTTTCTACCTGACCTTTTCTCATCCTGGGGACATGTACAC
TCTTCTTCGCCTTTGAGTGCCGCTACCTGGCTGTTTCTGCTGTCTCCTGCCATCCCTGTATTTGCTGCCATGCTCT
TCCTTTTCTCCATGGCTACACTGTTGAGGACCAGCTTCACTGACCCTGGAGTGATTCTCGGGCGCTACCATG
AAGCAGCTTTTCATAGAAATGGAGATAGAAGCTACCAATGGTGCGGTGCCCCAGGGCCAGCGACCAACCGCTCGTA
TCAAGAATTTCCAGATAAACAACCAGATTGTGAAACTGAAATACTGTTACACATGCAAGATCTTCCGGCCTCCCC
GGGCCTCCCATTTGCAGCATCTGTGACAACGTGTGTGGAGCGCTTCGACCATCACTGCCCCCTGGGTGGGGAATTGTG
TTGGAAAGAGGAACTACCGCTACTTCTACCTCTTCATCCTTTCTCTCCCTCCTCACAATCTATGTCTTCGCCT
TCAACATCGTCTATGTGGCCCTCAAATCTTTGAAATTTGGCTTCTTGGAGACATTGAAAGAAACTCCTGGAAGT
TTCTAGAAGTCCTCATTGCTTCTTTACACTCTGGTCCGTGCGTGGGACTGACTGGATTTTCTACTTTCTCTGTG
CTCTCAACCAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAGAATCGCGTCCAGAATCCCTACAGCC
ATGGCAATATTGTGAAGAAGTGTGTGAAGTGTGTGTGGCCCCCTTGCCCCCAGTGTGCTGGATCGAAGGGGTA
TTTTGGCACTGGAGGAAAGTGAAGTCGACCTCCAGTACTCAAGAGACCAGTAGCAGCCTCTTGCCACAGAGCC
CAGCCCCCACAGAACACCTGAACTCAAATGAGATGCCGGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAG
AGCCCCCAGAGCCACCACAGGAGGCAGCTGAAGCTGAGAAGTAGCCTATCTATGGAAGAGACTTTTGTGTTGTGTT
TAATTAGGGCTATGAGAGATTTAGGTGAGAAGTTAAACCTGAGACAGAGAGCAAGTAAGCTGTCCCTTTTAACT
GTTTTTCTTTGGTCTTTAGTCACCCAGTTGCACACTGGCATTTTCTTGCTGCAAGCTTTTTTAAATTTCTGAACT
CAAGGCAGTGGCAGAAGATGTGAGTCACTCTGATAACTGGAAAAATGGGTCTCTTGGGCCCTGGCACTGGTTCT
CCATGGCCTCAGCCACAGGGTCCCCTTGACCCCCCTCTCTCCCTCCAGATCCAGCCCTCCTGCTTGGGGTCA
TGGTCTCATTCTGGGGCTAAAAGTTTTTGAGACTGGCTCAAATCCTCCCAAGCTGCTGCACGTGCTGAGTCCAGA
GGCAGTCACAGAGACCTCTGGCCAGGGGATCCTAACTGGGTCTTGGGGTCTTCAGGACTGAAGAGGAGGGAGAG
TGGGGTCAGAAGATTCTCTGGCCACCAAGTGCCAGCATTGCCCAAAATCCTTTTAGGAATGGGACAGGTACCT
TCCACTTGTTGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGTTTTTCTTTTGACTCCTGCTCCCATTAGGAG
CAGGAATGGCAGTAATAAAAGTCTGCATTTTGGTCACTTTCTTTTCTCAGAGGAAGCCCGAGTGCTCACTTAAAC
ACTATCCCCTCAGACTCCCTGTGTGAGGCCTGCAGAGGCCCTGAATGCACAAATGGGAAACCAAGGCACAGAGAG
GCTCTCCTCTCCTCTCCTCTCCCCGATGTACCTCAAAAAAAAAAATGCTAACCAAGTTCTTCCATTAAAGCCT
CGGCTGAGTGAGGGAAGCCAGCACTGCTGCCCTCTCGGGTAACTCACCTAAGGCCTCGGCCACCTCTGGCT
ATGGTAACCACACTGGGGGCTTCTCCAAGCCCCGCTCTTCCAGCACTTCCACCGGCAGAGTCCCAGAGCCACTT
CACCTGGGGGTGGGCTGTGGCCCCCAGTCAGCTCTGCTCAGGACCTGCTCTATTTTCAGGGAAGAAGATTTATGT
ATTATATGTGGCTATATTTCTAGAGCACCTGTGTTTTCTCTTTCTAAGCCAGGGTCTGTCTGGATGACTTAT
GCGGTGGGGGAGTGTAACCGGAACTTTTTCATCTATTTGAAGGCGATTAAACTGTGTCTAATGCA

[illegible]

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCCTGTATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTTGAATTAGGTAT
TATAGGGATGGTGGGGTTGATTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCTTTCT
CCCACAGAGCNCTTCGACCATCACTGCCCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAACTA
CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCNCCTCACAATCTATGTCTTCGCCTTCA
ACATCGT

FIGURE 221

GTTGTGTCCTTCAGCAAAACAGTGGATTTAAATCTCCTTGACACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAGAAGAAAAAGAAGAAG
AAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC
GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATT
GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGTCTTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTTGTGCAAGTATCTCCCAAATTTGTAGAGATTTT
TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCCAAAGCGGTTGGCTTTTGTGAGTGAAGAC
GAATACTTGGAAATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTACCCGTGAACTATCCACCATACA
TTTCAGAAGCCAAGGGTACAGGTGTCCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCTCAGCAGAATTCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTTCCTCTCAAACTCATCTTCTTCAATGTCTCTG
AACATGACTATGGGAACTACACTTGCCTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTTGGTCCAGGCGCCGTGAGCGAGGTGAGCAACGGCACGTGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGACCTGCTTCTCAAATTTTGATGTGAGTGCC
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCACCAACACAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAAGAGTTTTAAAAAAGAAATTGAA
AATTGCCTTGACAGATATTTAGGTACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCACAGAGTGCCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAA
AAAAA

FIGURE 222

MKTIQPKMHNSISWAIFTGLAALCLFQGVPRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVLLSNTQTQYSIEIQNVVDVYDEGPYTCSVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGFFVSEDEYL
EIQGITREQSGDYECsasNDVAAPVVRVKVTVNYPPYISEAKGTGVPVGQKGTLQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNVTCVASNKLGHNTNASIML
FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLKLF

Signal peptide:

amino acids 1-28

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 223

GAAAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTA
TTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTTGTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

FIGURE 224

ATGGCTGGTGACGGCGGGGCCGGGCAGGGGACCGGGGCCGGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGA
ATCACCGCTGGCCCCGACTCCACCATGAAACGTGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAG
AAGGGGACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCT
GCCTGCTTCTGGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCATCCACAGCACCTGCCTTACA
GAGGCCTGCATTGAGTGGCTGGAAAAATCCTGGAGTCCCTGGACCGAGGGGTGAGCCCTGTGAGGACTTTTAC
CAGTTCTCTGTGGGGGCTGGATTGCGAGGAACCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGC
CTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAG
CAGAAGACACAGCGCTTCTACCTATCTTGCCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCCAGCCACTGAGA
GACCTCATTTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAGGACAACCTTTATGGAGGTGTTGAAG
GCAGTAGCAGGGACCTACAGGGCCACCCCATTTCTTACCCTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGC
AATGTTATCCAGGTGGACAGTCTGGGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAG
AAAGTGCTCACTGCCTATCTGGATTACATGGAGGAACCTGGGGATGCTGCTGGGTGGGCGGCCACCTCCACGAGG
GAGCAGATGCAGCAGGTGCTGGAGTTGGAGATACAGCTGGCCAACATCACAGTGCCCCAGGACCAGCGCGCGAC
GAGGAGAAGATCTACCACAAGATGAGCATTTCGGAGCTGCAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTC
CTGTCTTTCTGTGCTGTCACCATTTGAGTTGAGTGACTCTGAGCCTGTGGTGGTGTATGGGATGGATTATTTGCAG
CAGGTGTCAGAGCTCATCAACCGCACGGAACCAAGCATCCTGAACAATTACCTGATCTGGAACCTGGTGCAAAAG
ACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCCCTATATGGCACTAAGAAG
TCCTGTGTGCCGAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCTTGGCTTTGCTTTGGGGTCACTCTTC
GTGAAGGCCACGTTTGACCGGCAAAGCAAAGAAATTGCAGAGGGGATGATCAGCGAAATCCGGACCGCATTTGAG
GAGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAGAAAGCAGATGCCATCTAT
GATATGATTGGTTTCCCAGACTTTATCCTGGAGCCCAAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT
GAAGATTCTTTCTTCCAAAACATGTTGAATTTGTACAACTTCTCTGCCAAGGTTATGGCTGACCAGCTCCGCAAG
CCTCCCAGCCGAGACCACTGGAGCATGACCCCCAGACAGTGAATGCCCTACTACCTTCCAACCTAAGAATGAGATC
GTCTTCCCCGCTGGCATCCTGCAGGCCCTTCTATGCCCCGAACCAACCCCAAGGCCCTGAACCTTCGGTGGCATC
GGTGTGGTCACTGGGCCATGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTG
CGGCCCTGGTGGCAGAAAGTGCCTGGCAGCCTTCCGGAACCAACAGGCCTGCATGGAGGAACAGTACAATCAA
TACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACATTACTGACAACGGGGGGCTGAAG
GCTGCCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACC
AACCACCAGCTCTTCTTCGTGGGATTTGCCCAGGTGTGGTGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGG
CTGGTGACCGACCCCCACAGCCCTGCCCGCTTCCGCGTGTGGGCACTCTCTCCAACCTCCCGTGACTTCCTGCGG
CACTTTCGGCTGCCCTGTGGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACCTGGATCAGGGGA
GAAATGGCCAGCTGTCAACAGACCTGGGGCAGCTCTCCTGACAAAGCTGTTTGTCTTGGGTGGGAGGAAGCAA
ATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCCACAGGTGACATGAGTACAGACCCTCCTCAATCACCACATTG
TGCTCTGCTTTGGGGGTGCCCCCTGCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTCCGTGTCAACCT
GCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCC

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDPSH
STCLTEACIRVAGKILESIDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQNQA
ILKHLLNTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNITGPWDQDN
FMEVLKAVAGTYRATPFFTUYISADSKSSNSNVIQVDQSGFLPSRDYYLNRTANEKVLTA
LDYMEELGMLLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLLLETLYGTKKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTAFFEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMLNLYNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLPKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKKGEEQQLPVGLTNHQLFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPMNPQGQLCEVW

Type II Transmembrane domain:

amino acids 32-57

FIGURE 226

GCGCGGCCCTCCGCCCTCCGCACTCCCGCCTCCCTCCCTCCGCGCCCTCCTCCCTCCCTCCCTCC
 CAGCTGTCCCGTTCCGCTCATGCCGAGCCTCCCGGCCCGCGGCCCGCTGCTGCTCCTCGGGCTGCTGCTGCT
 CGGCTCCCGGCCCGGCCCGCGGCCCGGCCCGGCCCGCTGCTGCTGCTCCTCGGCTGCTGCTGCTGCT
 CGTTCCGGGAGCGGCAGGTAGGTGGGCGCCCGGGGAGGCGCGGGCGGGAGTCCGGCTCGGGCGAGTCA
 CAGCCCGGAGGGGGCGCGGGGCGCAGGTGGCTCGGCGCGGGCGGGCGGCCCGAGGGTGGGCGGGGGC
 GCGGTGCTGGGACCCGGGACCCGCGGGCAGCCCCCGGGGCGGCACACGGCGCGAGTGGGCGAGCGG
 CAAGCCCGTCCCGCAGGCTGCACCTTCGGCGGGAAGGTCTATGCCTTGGACGAGACGTGGCACCCG
 GGAGCCATTCCGGGTGATGCGCTGCGTGTGTGCGCCTGCGAGGCGCAGTGGGGTCCCGGTACCAG
 CAGGGTCACTGCAAGAACATCAAACCAAGATGCCAACCCCGCCCTGTGGGCGAGCCGCGCCAGCT
 CTGCTGCCAGACCTGCCCCAGGACTTCGTGGCGCTGCTGACAGGGCCGAGGTGCGAGGCGGTGGC
 AGTCTCGCTGCTGCGCTCTAGCCTCCGCTTCTCTATCTCCTACAGGCGGCTGGACCGCCCTACC
 CTCAGACTCCAATGGCAGTGTCTGTTGAGCACCTGCGAGCCCCACCAAGATGGCCTGGTCTGTGG
 GCGGGCAGTGCCTCGGTTGTCTCTGCGGCTCCTTAGGGCAGAACAGCTGCATGTGGCACTTGTG
 CCCTTCAGGGGAGGTCTGGGGGCTCTCATCCGGCACCGGGCCCTGTCCCCAGAGACCTTCAGT
 TCTAGAAGGCCCCCACCAGCAGGGCGTAGGGGGCATCACCTTGCTCACTCTCAGTGACACAGAGG
 TTTTTTGCTGCTCTTCCGAGGCTTGCAGGACTAACCCAGGTTCCCTTGAGGCTCCAGATTCTA
 GCTACTGCGAGAATTCAGGCCAATGTCTCAGCCAGGAACAGGCTTTGCTGAGGTGCTGCCAACCT
 CCAGGAGATGGACTGGCTGGTGTGGGGAGCTGCAGATGGCCCTGGAGTGGGCGAGGCGAGGCGT
 CAGTGGACACATTGCTGCCAGGAAGAGCTGCGACGTCCTGCAAAGTGTCTTTGTGGGGCTAATG
 AGTCCAAACGGGTGCTGCCGGCTCAGCCAGCCTCACTCTGCTAGGAAATGGCNCCTGATCCTC
 GGTAGGGACAACAGTGGTGGTGGCCATGACACTGGAAACCAAGCCTCAGCGGAGGGATCAGCC
 GTGCCACATGGCTGGCCTATCCTCCCTGCCCCAGGCGCTGGGTATCTGCCCTGGGCTGGGGT
 TCATATGCTGCTGCAGAAAGAGCTCTTCTGAACGTGGGCACCAAGGACTTCCAGACGGAGAG
 ACGTGGCTGCCCTGCCCTACTGTGGGGCATAGCGCCCGCTGCCCCGCTAGCAGGAGCCCTGGT
 CCCTGTGAAGAGCCAAGCAGCAGGGCACGCTGGCTTTCTTGGATACCCACTGTCACTGCACTAT
 GCTGGCTGGGCTTGGTGGCTCAGAACAGGCACTGTCACTGCCACCTCCTTGGGCTCCTGGAAC
 TCGGCGGCTGCTGAAGGATTCATGGCTCAGAGGCCAGGGTGTGGTGAAGGACCTGGAGCCG
 GCACCTGGCAAAAGGCATGGCTTCCCTGATGATCACCACCAAGGTAGCCCCAGAGGGGAGCT
 CTCTCCAGGTGCACATAGCCAACCAATGTGAGGTGGCGGACTGCGCCTGGAGGCGGCGGGG
 GCGGGCGCTGGGGGCTCCGGATACAGCCTCTGCTGCGCCGCTGTGGTGGCTGGTCTCCCGG
 CAAACCTGGTGGTCTTGGGCGGCCCGGAGACCCCAACACATGCTTCTTCGAGGGGCGAGC
 TCGCTGGGCGCCCACTACGACCCGCTCTGCTCACTCTGCACCTGCCAGAGACGAACGGTGA
 GGTGTGCCACCGCCAGCTGCCACACCCGCTGAGGCTCCCGACCAAGTGTGCTGCCCTGTTT
 TTTTGATGGTGAACGGAGCTGGCGGGCAGCGGTACGCGGTGGCACCCCGTTGTGCCCTTTG
 GTGTGCTGTCTGCACCTGCAAGCAGGGGGGCACTGGAGAGGTGCACTGTGAGAAGGTGCA
 CTGTGCCCAGCCTGTGCGTGTCAACCCACCGACTGCTGCAACAGTGTCCAGGTGAGGCGC
 GGACCCCATGCAGGCTGATGGGCCCCGGGGCTGCCGTTTGTGCTGGGCGAGTGGTTC
 CCCAGAGAGTCAAGCTGGCA CCCCTCAGTGCCCCGTTTGGAGAGATGAGCTGTATCACCTG
 GTGAGGTGGGTACTGGGAGCCTGGTCTGGAGTAGGGAGACCTTCCAGGGAGGTCCCTGA
 CTGTGTCCAGTGCCTCTGGGGACACTCAGTGTCTGCTCTGTCTGTACCAGGCGAGGGTGC
 GGGATGACTGTTCACTGCCACTGTCTGTGGCTCGGGGAAGGAGAGTGCATGCTGTTC
 GGCGGCGTAAGTGAGGGAGTCCAGGGTCAGCAGCTGTGAGTGGAGGGCTCACCTGCCTGT
 GGAAGGGAGCACTCACTGTGTGCAGGAACAGTGCAGCCTGCCTCACAAGTGCCATTCCAAT
 ACCTGGTGGAATTGTTATTTATGACCTTTTCTTTACAAATGAGATTTCTGAAGCTCAGAG
 ATGAAGGTCAACAGCTGTGTGCACTGACCTGTTTAGAAAATACTGGCCTTTCTGGGACCA
 TGCCCTGCCCTCTATGCCTCTCTGTGCCTCTCCACTCCCTCTCCCTCCTCCAACATTCCCT
 AGCAGCCCCAGAGACCAGAACTGATCCAGAGCTGGAGAAAGAGCCGAAGGCTTTAGGGAG
 AAGTGACCAAGAGGATGGGGCTGAGCTGGGGAAGGGGTGGCATCGAGGACCTTCTTGCA
 CCCAGTGCCTTTGCTCCTCTGTCTGCTCTACTCCCAACCCCACTACCTCTGGGAACCA
 GAGAGGACCTGGGCCAGACCGAGGTACAGCCACTCCAAGTCTGCCCCTGCCACCTCGGCT
 GCGCCCTTTCTTCTCTGTACATAATGTCACTGGCTTGTGGGATTTTAAATTTATCTTCACT
 GGCCCCGACACTCCACTCCTGTGCCCCCTGAGCTGAGCAGAGTCATTATTGGAGAGTTTGT
 ATTTCTTTTTCAGTCTTTGGGCATGAGGTGGCTCTTTGTGGCCAGGAACCTGAGTGGGGC
 GCNAGAGTAGGAGGTGAGAGAGAGGAGCTCTGACACTTGGGGAGCTGAAAGAGACCTGG
 CGTGGCNCNTTGGCTGGCATNCCTGGGTTCGCGAGAGGGGCTGGGGATGGTCTTGGAGAT
 AATTTAGGGAAGTAGAAGCAGGATTTTGACTCAAGTTTAGTTTCCACATCGCTGGCCTGT
 TTGAAGTTGCTCCAGAGAGAGAATCAAAGGTGTACCAGCCCCCTCTCTCCCTCCTTCCCTT
 TTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC

FIGURE 227

GGCCGAGCGGGGGTGCTGCGCGGCGGCCGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGA
CCGGGGCCGCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCCGAC
TCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC
TGCTGGCTGCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCA
TCCCACAGCACCTGCCTTACAGAGGCCTGCATTGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCCCTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATTGGA
GGAACCCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAAC
CAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC
AGCCACTGAGAGACCTCATTTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAG
GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCATTCTTCAC
CGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTACCTAGTGAACAAACTGCCCT
CCTTTCTTTCTTCTTTTCTTCTCCCTCCCTCCCTTTCTTCCCTTTTCTTCTCCTTCCCTCC
TCTTATTCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGGAATTG
GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTCAGTGTGATGGGTTTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAAGAGGGTCCTGAGGATTCAGGAGAGACTTATGGAGCC
AGCAAAGTCTTCTGAAGAGATTGCATTTGAGCCAGGTCCTGTAG

FIGURE 228

ATGCCTACTACCTTCCAAC TAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCCC
TTCTATGCCCCGAACCACCCCAAGGCCCTGAACTTCGGTGGCATCGGTGTGGTCATGGGCCA
TGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAG
TACAAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACAT
TGCTGACAACGGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATTT
GCCCAGGTGTGGTGTCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCC
CCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAAC TCCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTCTGGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGGCCAGCTGTCAACAGACCTGGGGCAGCTCTCCTGACAAAGCTGT
TTGCTCTTGGGTTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACA
GGTGACATGAGTACAGACCCTCCTCAATCACCACATTGTGCCTCTGCTTTGGGGGTGCCCTT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCCGTGTCACCCTGCCTGGAAGAG
GTCTGGGTGGGGAGGCCAGTTC CATAGGAAGGAGTCTGCCTCTTCTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGGCCTGCCGTGCCTGCCCCACTGTGACCCACAGGCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCGCACTTAGGGGTGGACTCAGCTCTGTC
TGGCTCACCCCTCACGGGCTACCCCCACCTCACCTGTGCTCCTTGTGCCACTGCTCCCAGTG
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGGCCTCTGAAAGCCTCCTGC
TGCCCCACTGTTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGGTACTGGTTCCTGT
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTTATTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT
TGGCCCTTATAGGACC

Figure 228

FIGURE 229

CCCACGCGTCCGAGCCGCCGAGAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGA
GGAGGGGAGGCAAAAACACCGAAAAACAAAAGAGAGAAACAACACCCAACTGGGGTGG
GGGGAAGAAAGAAAAGAAAAGAAACCCACCCACCCACCAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAATCCTGTGGCGCGCCGCTGGTTCCCGGGAAGACTCGCCAGCACCAGGGGG
TGGGGGAGTGCAGAGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGGATGGACATGATG
CTGTTGGTGCAGGGTGCTTGTGCTCGAACCAGTGGCTGGCGGCGGTGCTCCTCAGCCTGTG
CTGCCTGCTACCCTCCTGCCTCCCGGCTGGACAGAGTGTGGACTTCCCCTGGGCGGCCGTGG
ACAACATGATGGTCAGAAAAGGGGACACGGCGGTGCTTAGGTGTTATTTTGAAGATGGAGCT
TCAAAGGGTGCCTGGCTGAACCGGTCAAGTATTATTTTTCGCGGAGGTGATAAGTGGTCAGT
GGATCCTCGAGTTTCAATTTCAACATTGAATAAAAGGGACTACAGCCTCCAGATACAGAATG
TAGATGTGACAGATGATGGCCCATACACGTGTTCTGTTTCACTCAACATACACCCAGAACA
ATGCAGGTGCATCTAACTGTGCAAGTTCCTCCTAAGATATATGACATCTCAAATGATATGAC
CGTCAATGAAGGAACCAACGTCACTCTTACTTGTGTTGGCCACTGGGAAACCAGAGCCTTCCA
TTTCTTGGCGACACATCTCCCATCAGCAAAACCATTTGAAAATGGACAATATTTGGACATT
TATGGAATTACAAGGGACCAGGCTGGGGAATATGAATGCAGTGCAGGAAATGCTGTGTGCTT
CCAGATGTGAGGAAAGTAAAAGTTGTTGTCAACTTTGCTCCTACTATTAGGAAATTAAAT
CTGGCACCGTGACCCCCGGACGCGAGTGGCCTGATAAGATGTGAAGGTGCAGGTGTGCCGCT
CCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCTCTTCAATGGCCAACAAGGAATTATTAT
TCAAAATTTTAGCACAAAGATCCATTCTCACTGTTACCAACGTGACACAGGAGCACTTCGGCA
ATTATACCTGTGTGGCTGCCAACAAGCTAGGCACAACCAATGCGAGCCTGCCTCTTAACCT
CCAAGTACAGCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTTTCTCCTGCTGGTACCT
TGTGTTGACACTGTCCTCTTTCACCAGCATATTCTACCTGAAGAATGCCATTCTACAATAAA
TTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCTGAAAGTGCTGATGGCTGGATCCAATCT
GGTACAGTTTGTAAAAGCAGCGTGGGATATAATCAGCAGTGCTTACATGGGGATGATCGCC
TTCTGTAGAATTGCTCATTATGTAAATACTTTAATTCTACTCTTTTTTTGATTAGCTACATTA
CCTTGTGAAGCAGTACACATTGTCCTTTTTTTAAGACGTGAAAGCTCTGAAATTACTTTTAG
AGGATATTAATTGTGATTTCATGTTTGTAATCTACAACCTTTTCAAAGCATTTCAGTCATGGT
CTGCTAGGTTGCAGGCTGTAGTTTACAAAAACGAATATTGCAGTGAATATGTGATTCTTTAA
GGCTGCAATACAAGCATTCAGTTCCTGTGTTTCAATAAGAGTCAATCCACATTTACAAAGATG
CATTTTTTTTCTTTTTTGATAAAAAAGCAAATAATATTGCCTTCAGATTATTTCTTCAAATA
TAACACATATCTAGATTTTTCTGCTTGCATGATATTAGGTTTCAGGAATGAGCCTTGTAAT
ATAACTGGCTGTGCAGCTCTGCTTCTCTTTCCTGTAAGTTCAGCATGGGTGTGCCTTCATAC
AATAATATTTTTCTCTTTGTCTCCAATAATATAAAATGTTTTGCTAAATCTTACAATTTGA
AAGTAAAAATAAACCAGAGTGATCAAGTTAAACCATACACTATCTCTAAGTAACGAAGGAGC
TATTGGACTGTAAAAATCTCTTCTGCACTGACAATGGGGTTTGAGAATTTTGCCCCACACT
AACTCAGTTCTTGTGATGAGAGACAATTTAATAACAGTATAGTAAATATACCATATGATTTT
TTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGAAATCATTCCCTTTAAAATGACAGCACA
GTCCACTCAAAGGATTGCCTAGCAATACAGCATCTTTTCTTTTCACTAGTCCAAGCCAAAA
TTTTAAGATGATTTGTGAGAAAGGGCACAAAGTCCTATCACCTAATATTACAAGAGTTGGTA
AGCGCTCATCATTAATTTTATTTTGTGGCAGGTATTATGACAGTCGACCTGGAGGGTATGGA
TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACCAGGGTGGTTATGACCGCTACTC
AGGAGGAAATTACAGAGACAATTATGACAACCTGAAATGAGACATGCACATAATATAGATACA
CAAGGAATAATTTCTGATCCAGGATCGTCCTTCCAAATGGCTGTATTTATAAAGGTTTTTGG
AGCTGCACTGAAGCATCTTATTTTATAGTATATCAACCTTTTGTTTTTTAAATTGACCTGCCA
AGGTAGCTGAAGACCTTTTAGACAGTTCATCTTTTTTTTTTAAATTTTTTCTGCCTATTTAA
AGACAAATTATGGGACGTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLA AVL LSLCCLLPSCLPAGQSVDFPWA AVDNMMVRKGD TAVLR CYLED
GASKGAWLNRSSI IFAGGDKWSVDPRVSI STL NKRDSLQIQNV DVTDDGPYTCSVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTNVTLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYECSAENAVSFDPVRKVKV VVNFAPT IQEIKSGTVTPGRSGLIRCEGAGV
PPPAFEWYKGEKKLFNGQQGII IQNFSTRSILTVTNVTQEHFGNYTCVAANKLGTTNASLPL
NPPSTAQYGITGSADVLFSCWYLVLTLS SFTSIFYLKNAILQ

Important features of the protein:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325

Myelin P0 protein:

amino acids 92-121

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTTCTGCTGG
CTCCAGGACTTTGGCCATCTATAAAGCTTGGCAATGAGAAATAAGAAAATTCTCAAGGAGGA
CGAGCTCTTGAGTGAGACCCAACAAGCTGCTTTTCACCAAATTGCAATGGAGCCTTTCGAAA
TCAATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACCTTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGC
GCGGCTCCGGGTCTTGAGATGTATTTCTCAATGACACTCTGGCGGCTGAGGACAGCCCGT
CCTTCTCCTTGCTGCAGTCAGCACACCCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTG
CAAGTCCTGCAGGCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGT
AGACAACTTCACTCAGAACCCAGGGATGTTTCAGAATCAAAGGTGAACAAGGCGCCCCAGGTC
TTCAAGGTCACAAGGGGGCCATGGGCATGCCTGGTGGCCCTGGCCCCGCCGGGACCACCTGCT
GAGAAGGGAGCCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCCCTCGGGACCCCAAGG
CCCACCGGGAGTCAAGGGAGAGGCGGGCCTCCAAGGACCCAGGGTGCTCCAGGGAAGCAAG
GAGCCACTGGCACCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATT
GGCCCCAAAAGGGGAACTGGAAC TAAGGGAGAGAAAGGAGACCTGGGTCTCCAGGAAGCAA
AGGGGACAGGGGCATGAAAGGAGATGCAGGGGT CATGGGGCCTCCTGGAGCCAGGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCCACCAGGTTTGGCTGGTTTTCTGAGCTAAAGGAGAT
CAAGGACAACCTGGACTGCAGGGTGTTCCGGGGCCCTCCTGGTGCAGTGGGACACCCAGGTGC
CAAGGGTGAGCCTGGCAGTGCTGGCTCCCCTGGGCGAGCAGGACTTCCAGGGAGCCCCGGGA
GTCCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGA
AGAAAAGGAGAATCAGGAGTTCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCCAGG
GCTGGCAGGTCCCAAGGGAGCCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAAG
GATCTTCTGGGGAGCAAGGAGTAAAGGGAGAAAAAGGTGAAAGAGGTGAAAACCTCAGTGTCC
GTCAGGATTGTCTGGCAGTAGTAACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGG
GACAATTTGCGATGACGAGTGGCAAAATTCTGATGCCATTGTCTTCTGCCGCATGCTGGGTT
ACTCCAAAGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAAT
GTTTCAAGTGTCTGGGGCACGGAGAGTACCCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCA
TGACTGCAGCCACGAGGAGGACGCAGGCGTGGAGTGCAGCGTCTGAACCCGGAAACCCTTTCA
CTTCTCTGCTCCCGAGGTGTCTCGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT
TCCCTGGGGACAACCTGAGCAGCCTCTGGAGAGGGGCCATTAATAAAGCTCAACATCATTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKEDELLSETQQAAPHQIAMEPFEINVPKPKRRNGVNFSLAVVVIYLILLTAGAGL
LVVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLLQSAHPGEHLAQGASRLQVLQAQLTWVR
VSHEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRD
GATGPSGPPQPGVKGEGAGLQGPQGAPGKQGATGTPGPQGEKGSKGDGGLIGPKGETGTKGE
KGDGLPGSKGDRGMKGDAGVMGPPGAQGSKGDFGRPGPPGLAGFPAGKGDQGPGLQGVPG
PPGAVGHHPGAKGEPGSAGSPGRAGLPSPGSPGATGLKGSKGDGTGLQGGQGRKGESGVPGPA
GVKGEQGSPLAGPKGAPGQAGQKGDQGVKGSSGEQGVKGEKGERGENSVSVRIVGSSNRGR
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWLDDNVQCRGTESTLW
SCTKNSWGHHDCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

Clq domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 234

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758

<subunit 1 of 1, 98 aa, 1 stop

<MW: 11081, pI: 6.68, NX(S/T): 1

MKLMVLVFTIGLTLLLGVQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG

KGCEMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

FIGURE 235

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCAGGTCTGGAGCGAATTCAGCCTGCAGGG
CTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTTACCCCGCCGTGGTGGTTGGAGGGC
GCGCAGTAGAGCAGCAGCACAGGCGCGGGTCCCGGGAGGCCGGCTCTGCTCGCGCCGAGATG
TGGAATCTCCTTCACGAAACCGACTCGGCTGTGGCCACCGCGCGCCCGCGCTGGCTGTG
CGCTGGGGCGCTGGTGCTGGCGGGTGGCTTCTTTCTCCTCGGCTTCCTCTTCGGGTGGTTTA
TAAAATCCTCCAATGAAGCTACTAACATTACTCCAAAGCATAATATGAAAGCATTTTTGGAT
GAATTGAAAGCTGAGAACATCAAGAAGTTCTTACATAATTTTACACAGATACCACATTTAGC
AGGAACAGAAACAAAACCTTTAGCTTGCAAAGCAAATTCATCCAGTGGAAGAATTTGGCC
TGGATTCTGTTGAGCTAGCTCATTATGATGTCTGTCTACCCAAATAAGACTCATCCC
AACTACATCTCAATAATTAATGAAGATGGAAATGAGATTTTCAACACATCATTATTTGAACC
ACCTCCTCCAGGATATGAAAATGTTTCGGATATTGTACCACCTTTAGTGCTTCTCTCCTC
AAGGAATGCCAGAGGGCGATCTAGTGATGTTAACTATGCACGAACTGAAGACTTCTTTAAA
TTGGAACGGGACATGAAAATCAATTGCTCTGGGAAAATTGTAATTGCCAGATATGGGAAAGT
TTTCAGAGGAAATAAGGTTAAAAATGCCAGCTGGCAGGGGCCAAAGGAGTCATTCTCTACT
CCGACCCTGCTGACTACTTTGCTCCTGGGGTGAAGTCCTATCCAGACGGTTGGAATCTTCCT
GGAGGTGGTGTCCAGCGTGGAAATATCCTAAATCTGAATGGTGCAGGAGACCTCTCACACC
AGGTTACCCAGCAAATGAATATGCTTATAGGCGTGGAATTGCAGAGGCTGTTGGTCTTCCAA
GTATTCCTGTTTCATCCAATTGGATACTATGATGCACAGAAGCTCCTAGAAAAAATGGGTGGC
TCAGCACCACCAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCCTACAATGTTGGACCTGG
CTTTACTGGAACTTTTCTACACAAAAGTCAAGATGCACATCCACTCTACCAATGAAGTGA
CGAGAATTTACAATGTGATAGGTACTCTCAGAGGAGCAGTGGAACCAGACAGATATGTCATT
CTGGGAGGTCACCGGACTCATGGGTGTTTGGTGGTATTGACCCTCAGAGTGGAGCAGCTGT
TGTTTCATGAAATTGTGAGGAGCTTTGGAACACTGAAAAGGAAGGGTGGAGACCTAGAAGAA
CAATTTTGTGTTGCAAGCTGGGATGCAGAAGAATTGGTCTTCTTGTTTCTACTGAGTGGGCA
GAGGAGAATTCAGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTAT
AGAAGGAACTACACTCTGAGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACACAACC
TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAAGGCAAATCTCTTTATGAAAGTTGG
ACTAAAAAAGTCCTTCCCCAGAGTTTCAGTGGCATGCCCAGGATAAGCAAATTGGGATCTGG
AAATGATTTTGAGGTGTTCTTCCAACGACTTGGAAATGCTTCAGGCAGAGCACGGTATACTA
AAAATTGGGAAACAAACAAATTCAGCGGCTATCCACTGTATCACAGTGTCTATGAAACATAT
GAGTTGGTGGAAAAGTTTTATGATCCAATGTTTAAATATCACCTCACTGTGGCCAGGTTTCG
AGGAGGGATGGTGTGTTGAGCTAGCCAATTCATAGTGCTCCCTTTTGATTGTCGAGATTATG
CTGTAGTTTTAAGAAAGTATGCTGACAAAATCTACAGTATTTCTATGAAACATCCACAGGAA
ATGAAGACATACAGTGATCATTTGATTCACTTTTTTCTGCAGTAAAGAATTTTACAGAAAT
TGCTTCCAAGTTCAGTGAGAGACTCCAGGACTTTGACAAAAGCAACCCAATAGTATTAAGAA
TGATGAATGATCAACTCATGTTTCTGGAAAGAGCATTATTTGATCCATTAGGGTTACCAGAC
AGGCCTTTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC
ATTCCCAGGAATTTATGATGCTCTGTTTGATATTGAAAGCAAAGTGGACCTTCCAAGGCCT
GGGGAGAAGTGAAGAGACAGATTTATGTTGCAGCCTTCACAGTGCAGGCAGCTGCAGAGACT
TTGAGTGAAGTAGCCTAAGAGGATTTTTTAGAGAATCCGTATTGAATTTGTGTGGTATGTCA
CTCAGAAAGAATCGTAATGGGTATATTGATAAATTTTAAAATTGGTATATTGAAATAAAGT
TGAATATTATATATAA

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756

><subunit 1 of 1, 750 aa, 1 stop

><MW: 84305, pI: 6.93, NX(S/T): 10

MWNLLHETDSAVATARRPRWLCAGALVLAGGFLLGFLFGWFIKSSNEATNITPKHNMKAFL
DELKAENIKKFLHNFTQIPHLAGTEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH
PNYISIINEDGNEIFNTSLFEPPPPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINCSGKIVIRYGVFRGNKVKNQAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQQRGNILNLNGAGDPLTPGYPANAYARRGIAEAVGLPSIPVHPIGYYDAQKLEKMG
GSAPPDSSWRGSLKVPYNVGPFGFTGNFSTQKVKMHISTNEVTRIYNVIGTLRGAVEPDYV
ILGGHRDSWVFGGIDPQSGAAVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLQLQERGVAYINADSSIEGNYTLRVDCTPLMYSLVHNLTKELKSPDEGFEGKSLYES
WTKKSPSPEFSGMPRISKLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET
YELVEKFYDPMFKYHLTVAQVRGGMVFEANISIVLPFDCRDYAVVLRKYADKIYSISMKHPQ
EMKTYSVSFDSLFSVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAFIDPLGLP
DRPFYRHVIYAPSSHKNKYAGESFPGIYDALFDIESKVDPSKAWGEVKRQIYVAAFTVQAAAE
TLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713